DOCUMENT RESUME

ED 065 089 HE 003 249

AUTHOR Calista, Donald J.
TITLE College Students' Instructional Expectations and Evaluations.
PUB DATE Jul 72
NOTE 134p.

EDRS PRICE MF-$0.65 HC-$6.58
DESCRIPTORS *Effective Teaching; *Evaluation Methods; *Higher Education; *Relevance (Education); Summative Evaluation; *Teacher Evaluation; Teacher Improvement; Teaching Quality

ABSTRACT

Typical end-of-course faculty ratings were questioned for their inability to measure actual classroom interaction. Extending the concept of these evaluations to include the student instructional expectations dimension, the study proposed that the classroom experience be related to the process and systems approaches, more dependent upon monitoring and feedback. As a working model for classroom role structure, it employed Stogdill's group achievement framework to distinguish between those expectations associated with institutional dynamics--called situational--and those related to the classroom--called contextual. Instruments were administered at the opening--and close--of a semester to 209 students in 8 social science classes in 3 colleges. Developed from previous Critical Incident and factor analysis techniques, the 20 instrument items employed modified Semantic Differential scales. Additionally, the study examined relationships between expectations and the AVL (Allport-Vernon-Lindzey) Value Scale. The substantive hypotheses--for each class and demographic variable--found statistically significant differences between expectations and evaluations, even for high evaluations. If these evaluations demonstrated the direction of expectations, but failing to indicate their intensity, the study concluded that extrapolations made solely from evaluations were of questionable value in measuring the range of classroom dynamics. AVL intercorrelations did not produce any significant patterns, however. (Author)

ED 065089

FILMED FROM BEST AVAILABLE COPY

COLLEGE STUDENTS' INSTRUCTIONAL EXPECTATIONS AND EVALUATIONS

Donald J. Calista

Chairman, Behavioral and Social Sciences Department

Bennett College, New York

Millbrook 12545

HE 003 249

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ACKNOWLEDGEMENTS

This paper developed with the encouragement, counsel, and understanding of a number of individuals to whom I am most indebted.

First, and foremost, Dr. J. William Nystrom, Dean of the College in Bennett College, who gave freely of his time and critical abilities, and who quided me with wisdom and kindness. Discussions with his former doctoral colleague at Yale University, Professor Max Weiner, now Executive Officer of Doctoral Programs in Education in The City University of New York's Center for Advanced Study in Education, regarding experimental design and measurement technique proved most productive. To their wives a special thanks for allowing me to share their husbands' time. Professor Weiner's generosity enabled me to receive computer programming assistance with Mr. Charles A. Dyer, to whom I owe many hours saved at a desk calculator. Early advice on instrumentation from Dr. Anthony T. Santonicola, Director of Counseling Services in the University of Hartford and Dr. David Vitrogan of Yeshiva University was most valuable. Though not directly involved with the study, President Donald A. Eldridge of Bennett College, helped make the funds available to pursue my research interests, and aided me in many other ways during the past seven years.

The study could have never gotten off the ground without the cooperation of the participating instructors who graciously offered some of their classes' time. And to the two-hundred and nine nameless students, I trust their labors were worthwhile. My departmental colleagues in Bennett College, especially, Professor Joel Diemond, gave openly of their ideas and criticism. The staff of the College's Aldrich Library worked hard to round-up books and journal

articles quickly through the Mid-Hudson Interlibrary Loan System. To President Charles L. Palermo, of Laurence University, a special gratitude for his forbearance and levelheadedness throughout the research project. Mrs. Frieda Malament typed the final version efficiently and patiently.

A debt to one's family underscores the cooperative nature of a research venture. To my parents I owe lessons learned in patience. My son, Donald, and wife, RoseMarie, contributed much to maintaining equilibrium in a pressured situation. How self-indulgent it is not to share authorship.

D. J. C.
July 1972

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS ii

LIST OF TABLES vi

INTRODUCTION 1

Chapter

I. THE PROBLEM OF RELATING EXPECTATIONS TO EVALUATIONS 6

The Student As Evaluator 7

Developing Traits of the Ideal Professor 12

Reformulating the Evaluations Framework . . . 23

II. A FRAMEWORK FOR STUDENT ROLE EXPECTATIONS OF INSTRUCTORS 27

The Classroom as a Consensual Experience . . . 29

Consensual Expectations and Group Dynamics Theory 36

Relating Classroom and College Expectations . . 45

III. DESIGN, RESEARCH PROCEDURES AND HYPOTHESES 50

Study Design and Procedures 50

Developing an Instrument 53

Scaling Technique Employed 59

Hypotheses and Questions Considered 64

TABLE OF CONTENTS (continued)

Chapter		Page
IV.	PRESENTATION OF FINDINGS	68
	Hypotheses Considered	69
	Questions Considered	93
V.	DISCUSSION, IMPLICATIONS, AND RECOMMENDATIONS	99
	Discussion of Findings	99
	Implications	102
	Recommendations for Further Study	104
	Afterword	106
BIBLIOGRAPHY		109
APPENDIX A	Expectations Instrument	117
APPENDIX B	Evaluations Instrument	122

LIST OF TABLES

Table		Page
1.	Means and Standard Deviations for Expectations and Evaluations Compared by CI Category for the Four Classes in Alpha College and College Total	70
2.	Means and Standard Deviations for Expectations and Evaluations Compared by CI Category for the Two Classes in Beta College and College Total	71
3.	Means and Standard Deviations for Expectations and Evaluations Compared by CI Category for the Two Classes in Gamma College and College Total	72
4.	Comparison of Rank Differences by CI Expectations Category for Alpha, Beta and Gamma Colleges	74
5.	Comparison by Rank Differences of CI Expectations Category with CI Evaluations Category for Alpha, Beta and Gamma Colleges	75
6.	Means and Standard Deviations of Student Responses to Part II, Item nine on the Expectations and Evaluations Instruments for Alpha, Beta and Gamma Colleges by Sex	77
7.	Means and Standard Deviation Comparison of Freshmen with Sophomores by CI Expectations Category within Alpha, Beta and Gamma Colleges	78
8.	Means and Standard Deviation Comparison of Freshmen with Sophomores by CI Evaluations Category within Alpha, Beta and Gamma Colleges	80

LIST OF TABLES (continued)

Table — Page

9. Means and Standard Deviations by CI Expectations and Evaluations Category compared by Year for Alpha, Beta and Gamma Colleges 82

10. Means and Standard Deviations by Expectations and Evaluations CI Category Comparing Public with Private Secondary School Graduates According to Alpha, Beta and Gamma Colleges 84

11. Means and Standard Deviations by Expectations and Evaluations CI Category Comparing Students by Size of Residence City According to Alpha, Beta and Gamma Colleges 85

12. Means and Standard Deviations by Expectations and Evaluations CI Category Comparing Students by Social Science Majors with Non-Social Science Majors According to Alpha, Beta and Gamma Colleges 86

13. Means and Standard Deviations by Expectations and Evaluations CI Category Comparing Students by Previous Instructor Rating Experience According to Alpha, Beta and Gamma Colleges 88

14. Means and Standard Deviations by Expectations and Evaluations CI Category Comparing Students by Elective with Required Course Status According to Alpha, Beta and Gamma Colleges 89

15. Means and Standard Deviations by Expectations and Evaluations CI Category Comparing Students by Secondary School Class Size According to Alpha, Beta and Gamma Colleges 90

LIST OF TABLES (continued)

Table — Page

16. Means and Standard Deviations by Expectations and Evaluations CI Category Comparing Students by Sex According to Beta and Gamma Colleges 91

17. Means and Standard Deviations for AVL Scales on Pretest and Posttest by Sex and Year in Alpha and Beta Colleges 94

18. Means and Standard Deviations for AVL Scales on Pretest and Posttest by Sex and Year in Gamma College 95

19. Intercorrelation Matrix of Pretest AVL Values with CI Expectations Categories by Sex for Alpha, Beta and Gamma Colleges (decimal omitted) 97

20. Intercorrelation Matrix of Posttest AVL Values with CI Evaluations Categories by Sex for Alpha, Beta and Gamma Colleges (decimal omitted) 98

INTRODUCTION

This research project began largely by chance and circumstance. About three years ago a group of students active in winning acceptance for a student academic affairs committee at our college, asked me to help construct an end-of-course evaluation instrument. In the few hours allotted me by those busy and involved students, an instrument was prepared. Its subsequent use failed to stir up the anticipated anxieties; some faculty even welcomed it. As I shared the usual curiosity regarding faculty evaluation, I began gathering information about the field and tried various home-grown experiments in self-evaluation. A search through the literature, however, proved disheartening. Despite the growing enthusiasm of students for faculty evaluation, most investigators seemed to lack concern with integrating their findings into a comprehensive conceptual framework. Moreover, some appeared to be unaware or unconcerned with growing student demand that effective teaching and effective learning no longer be treated as separate issues.

In developing a research prospective, as described in this thesis, one basic question kept recurring: did course evaluations begin at the right juncture in the student-teacher relationship? Clearly, any realistic assessment of higher education, including evaluations research,

demanded the inclusion of both teacher and student domains.

In the broader sense this thesis developed at a time when critical perspectives regarding the student-teacher relationship are emerging. If this relationship served as a basic education unit, recent works maintained that higher education must reacquaint itself with its importance. The work of Joseph Katz was most influential; he challenged investigators in undergraduate education to rethink their traditional models by accepting the ideal of the fullest possible development of the individual student. Katz and his associates at Stanford University went beyond Sanford's earlier notion that curriculum must be a function of individual personality development, by concluding that this development must be reflected in all aspects of college experience.[1]

On a more polemical note, Harold Taylor voiced concern in sweeping terms:[2]

> Learning and teaching . . . have to do with the totality of human conduct, in which the conduct of affairs of the mind is by turns, political, social, public, private, intellectual, emotional, external, internal, and, in the last analysis, personal. Otherwise conduct has not meaning, the human act is stripped of motivation, empty of content, lacking in truth.

In reviewing recent works in higher education, Austin declared that although not all would completely support Taylor, educators and

---

[1]Joseph Katz, et al., No Time For Youth (San Francisco: Jossey-Bass, Inc., 1968), p. 160.

[2]Harold Taylor, Students Without Teachers: The Crisis in the University (New York: McGraw-Hill Book Company, 1969), p. 321.

researchers agreed that "a relevant curriculum must give attention both to the growth of the person and the building of a better society, neither of which can be sought in isolation from the other."[3]

Accepting this charge, Chickering advanced seven basic developmental tasks associated with college experience.[4] They were: achieving competence, managing emotions, becoming autonomous establishing identity, freeing interpersonal relations, clarifying purposes, and developing integrity. Although Chickering offered suggestions to implement these developmental objectives, notably by increasing the meaningful interaction of instructor and student, similar expressions characterizing the classroom as a community were proposed by Schwab.[5] He envisioned the classroom relationship essentially as a collaborative one, one that will be proposed by this thesis. Schwab stated that if the classroom presented an obvious "visibility of roles," there must also be "an exchange of roles where exchange is proper--so that students discover what it is like to be a professor and professors rediscover what it is like to be a student. It is a human society--not a society of disembodied minds; when professors confront students in the seminar room, they do so as persons who can be annoyed or irritated by students,

---

[3]C. Grey Austin, "The Year's Books," Journal of Higher Education, 40 (November, 1969), 589-592.

[4]Arthur W. Chickering, Education and Identity (San Francisco, Jossey-Bass, Inc., 1969), pp. 9-19.

[5]Joseph J. Schwab, College Curriculum and Student Protest (Chicago: University of Chicago Press, 1969).

who can be impatient, tired, distracted, who are concerned for students as well as with them--and by this visible freedom, enable students to be similarly free."[6] His emphasis was not on role prescription, but on role sharing. As Chickering's and Schwab's conceptions reflected changing pedagogical and organizational patterns in higher education, it became imperative not only to learn more about the behavior of student evaluations, but student expectations of instructors as well.

If the concepts of expectations and evaluations received wide circulation in the behavioral and social sciences, both seemed to lack any clear-cut or consistent definitions. For the purposes of this study, however, working definitions may be borrowed from ordinary language. Although expectations related to instruction will be discussed later, it can be defined here as the degree of probable occurrence of something believed to be desirable. Evaluations would be the summing up or measuring of the relative worth of actual occurrences. As used here, evaluations determined instructor effectiveness by employing rating instruments.

This study was organized in the following way: the first chapter reviewed how previous evaluations studies contributed to the expectations-evaluations framework developed here. The second chapter explored the nature of instructional expectations. Chapter three outlined

---

[6]Joseph J. Schwab, College Curriculum and Student Protest (Chicago: University of Chicago Press, 1969), p. 230.

research procedures and offered several hypotheses. The fourth chapter presented the findings. And chapter five discussed conclusions and implications of the study.

CHAPTER I

THE PROBLEM OF RELATING EXPECTATIONS TO EVALUATIONS

The controversy over student evaluations of their instructors does not suffer from the lack of empirical evidence. A recent review of the literature, with its bibliography of almost 300 recent research items has indicated that student evaluations could become more acceptable as important aspects of college education.[1] Yet the state of current research does not mask the continued debate regarding student evaluators. In a recent panel discussion of educators a panelist asked: "The question, ladies and gentlemen, is not why evaluation, not whether evaluation, but how evaluation is to be accomplished?"[2] This only draws attention to the basic questions addressed by previous student evaluations research.[3] The question could contribute, nonetheless, to a clearer

---

[1]Kenneth E. Eble, The Recognition and Evaluation of Teaching, Report of the Project to Improve College Teaching, Salt Lake City, Utah, November, 1970 (Salt Lake City, Utah: Project to Improve College Teaching, 1970), pp. 101-111.

[2]Walter Schoen, "Faculty Evaluation, Pro and Con: The Hypocrites," Report of the Twenty-Second Annual Conference, 1969 (Cazenova, New York: New York State Association of Junior Colleges, 1969), p. 34.

[3]Ibid., pp. 34-41.

understanding of the evaluative process by raising a prior question; namely, what is being measured by students in these instructor ratings? The argument of this paper is that the ratings measure what students expect of the instructor. Hence, this paper focuses on college student role expectations of instructors.

The Student as Evaluator

In order to determine how this perspective developed, the rationale and findings of the evaluations field itself must be examined. The most prominent form of evaluations research focused on end-of-course student ratings. Characteristically, these studies began with the premise that "as higher education is organized and operated, students are pretty much the only ones who observe and are in a position to judge the teacher's effectiveness."[4] A recent writer added that, "at present the only regular observers of the teacher in action are his students."[5] Generally, researchers correlated one intervening student variable such as sex or grades

---

[4]H. H. Remmers, "On Students' Perceptions of Teachers' Effectiveness," The Appraisal of Teaching In Large Universities, ed. W. J. McKeachie (Ann Arbor: University of Michigan Press, 1959), 20.

[5]Paul H. Owen, "Some Dimensions of College Teaching: An Exploratory Study Using Critical Incidents and Factor Analysis of Student Ratings" (unpublished Doctoral dissertation, University of Houston, 1967), p. 11.

with end-of-term instructor ratings. Presumably, this would uncover whether students were unduly prejudiced in their ratings. A leading figure in the field since the 1920's, H. H. Remmers, who conducted a number of rating studies at Purdue University, summarized the field's major research findings up to 1959. Those related directly to college teaching were:[6]

1. Grades of students are not in general closely related to the ratings of the teacher.
2. There is evidence showing that there is little, if any, relationship existing between student ratings of teachers and the judged difficulty of the course.
3. In a given institution there exists wide and important departmental differences in effectiveness of teaching as judged by student opinion.
4. The sex of the student rater bears little or no relationship to the ratings of teachers.
5. While the effect on student ratings of a generalized attitude (the 'halo effect') toward the teacher has not been isolated, it apparently does not exist to an extent sufficient to invalidate the ratings of separate aspects of teaching methods and of the course. Evidence indicates that students discriminate reliably for different aspects of teacher's personality and of the course and between different instructors and courses.
6. Teachers with less than five years experience tend to be rated lower than teachers with more than eight years experience.
7. Mature alumni of ten years' standing agree substantially with on-campus students in their evaluation of teachers.
8. The year in school of the rater has no effect on the ratings given except that ratings by graduate students tend to be higher than those by undergraduates.

---

[6]H. H. Remmers, "On Students' Perceptions of Teachers' Effectiveness," The Appraisal of Teaching In Large Universities, ed. W. J. McKeachie (Ann Arbor: University of Michigan Press, 1959), 21-22.

He cited two other findings not yet fully studied in the college setting, that "a considerable number of those who have used student rating believe this procedure is useful for facilitating the educational process," and, knowledge of student opinions and attitudes leads to improvement of the teacher's personality and educational procedures."[7] In conclusion, he noted: "No research has been published invalidating the use of student opinion as one criterion of teacher effectiveness."[8] In fact, in proposing a rating instrument for general use, McKeachie restated most of the earlier findings and reemphasized the reliability of student opinion in evaluation.[9]

Updating these conclusions, investigations in the last ten years generally followed the single-variable and end-of-course rating pattern. Widespread evidence reconfirmed that neither age, school-year level, grade-point average, sex or course grades

---

[7]H. H. Remmers, "On Students' Perceptions of Teachers' Effectiveness," The Appraisal of Teaching In Large Universities, ed. W. J. McKeachie (Ann Arbor: University of Michigan Press, 1959), 21.

[8]Ibid., p. 22.

[9]W. J. McKeachie, "Student Ratings of Faculty," AAUP Bulletin, 55 (December, 1969), 439-444.

adversely affected ratings.[10] One study at the California State College at Long Beach pointed out that even when students were actually cued to previous ratings of an instructor, they did not passively record their predecessors view.[11] As anticipated, a number of studies still focused on the influence expected grades had on ratings. A study done at the University of California at Berkeley supported the proposition that the rating scores will not be strongly biased by the leniency of the instructor in assigning grades.[12] While another, at the University of Washington extended these conclusions: "High ratings cannot be 'bought' by giving high grades, nor are they lost by giving low grades. Both when judging their instructor's over-all value as a teacher and when rating his skill in specific respects, such as clarity of presentation and development of interest the students rarely, if ever, were influenced by the grades

---

[10]Nicholas F. Rayder, "College Student Ratings of Instructors," Journal of Experimental Education, 37 (Winter, 1968), 76-81, and William E. Coffman, "Determining Students' Concepts of Effective Teaching from Their Ratings of Instructors," Journal of Educational Psychology, 45 (May, 1954), 277-286.

[11]James N. McClelland, "The Effect of Student Evaluations of College Instruction Upon Subsequent Evaluations," California Journal of Educational Research, 31 (March, 1970), 88-95.

[12]C. M. Garevick and H. D. Carter, "Instructors Ratings and Expected Grades," California Journal of Educational Research, 13 (November, 1962), 218-221.

which they had received from that teacher."[13] Seeking to determine the reliability of student ratings when different instructions were used on the same class, a study done at Brooklyn College reported "that students' ratings of teachers . . . are not greatly influenced by the different administrations of the questionnaires."[14] Moreover, no significant pattern relating authoritarianism in students with either high or low ratings was found at a "large southern woman's college."[15] An earlier study demonstrated that a panel of instructors plus administrators along with a random sample of students did not differ significantly in their assessment of the effectiveness of actual classroom situations.[16] A recent study at Clemson University re-confirmed reports that "personal qualities of the teacher, sex of the student, grades in the class, and overall GPR [grade-point rating]

---

[13]Virginia W. Voeks and G. M. French, "Are Student-Ratings of Teachers Affected by Grades?" Journal of Higher Education, 31 (June, 1960), 330-334.

[14]Manuel Cynamon and Shirley U. Wedeen, "Emotional Factors in Reliability of Student Ratings of Teachers," Journal of Educational Research, 51 (April, 1958), 629-632.

[15]Ann C. Maney, "Authoritarianism Dimension in Student Evaluations of Faculty," Journal of Educational Sociology, 32 (January, 1959), 226-231.

[16]Ruth E. Eckert, "Ways of Evaluating Teaching," School and Society, 71 (February 4, 1950), 65-69.

did not influence student assessment of teacher performance."[17]

If this evidence established that "college students appear to have greater objectivity and less superficial value systems than we have realized,"[18] another conclusion should be that the more researchers tried to uncover intervening variables influencing--that is, prejudicing--student choices, the more the findings discounted them. Few empirical studies, moreover, offered any substantial evidence disputing the concept of student end-of-term evaluations.

Developing Traits of the Ideal Professor

With the cumulative results of ratings studies giving impressive testimony to the judiciousness of students as evaluators, the second notable trend of evaluations research identified the image of the ideal professor. Concerned with eliciting general or overall characteristics of an ideal professor, these studies utilized both closed and open-ended instrument techniques. More recently, some studies examined whether differences associated with a student's major field, grade-point average, sex, year in school, as in the ratings studies,

---

[17]Bernard Caffrey, "Lack of Bias in Student Evaluations of Teachers," Proceedings of the 77th Annual Convention of the American Psychological Association, Vol. IV, Pt. II, 1969 (Washington, D. C.: American Psychological Association, 1969), pp. 641-642.

[18]Virginia W. Voeks and G. M. French, "Are Student-Ratings of Teachers Affected by Grades?" Journal of Higher Education, 31 (June, 1960), 330-334.

influenced perceptions of the ideal professor. Despite certain differences a general pattern of the ideal professor emerged.

The earliest, and still most systematic attempt was a study of an entire college undertaken by Riley and his associates at Brooklyn College in 1950.[19] Deriving a list of some twenty items from a previously administered open-ended questionnaire, students were asked to rank them in order of importance as teaching attributes. The ten items most often selected were the following:[20]

1. Knowledge of the subject.....88%
2. Attitude toward the subject.....78%
3. Organization of subject matter.....75%
4. Attitude toward students.....72%
5. Personality of instructor.....72%
6. Speaking ability.....72%
7. Ability to explain.....71%
8. Tolerance to disagreement.....71%
9. Fairness in examinations.....60%
10. Encouragement to thinking.....55%

A similar breakdown resulted from other studies employing open-ended techniques. Asking students to write a theme relating "the most effective professor they knew," Crawford and Bradshaw[21]

---

[19]John W. Riley, Jr., Bryce F. Ryan, and Marcia Lifshitz, The Student Looks at His Teacher (New Brunswick, New Jersey: Rutgers University Press, 1950). The sample numbered more than 6,000 students.

[20]Ibid., p. 82.

[21]P. L. Crawford and H. L. Bradshaw, "Perception of Characteristics of Effective University Teachers: A Scaling Analysis," Educational and Psychological Measurement, 28 (Winter, 1968), 1079-1085.

of Ohio University reported a list of behaviors unusually similar to Riley's rank order:[22]

1. Has a thorough knowledge of subject matter plus substantial knowledge in other fields.
2. Lectures are well planned and organized.
3. Is enthusiastic, energetic and has a lively interest in teaching.
4. Is student-oriented; willing to help students outside of class.
5. Encourages student participation in class by questions and discussion.
6. Relates class material to other fields.
7. Speaks clearly and distinctly.
8. Lacks defensive attitudes and prejudices.
9. Defines clearly the basis for evaluation of students' performances.
10. Uses a variety of teaching devices, demonstrations, charts.

In a study at Western Washington State College, Gadzella, likewise, examined the most and least prominent characteristics of college instructors. The five most important ones were: "knowledge of subject, interest in subject, flexibility, organization of daily and course preparations, and presentation of course materials." And the five least noted were: "publisher-writer, community participator, researcher, appearance and punctuality."[23] Quick and Wolfe considered the qualities students at the University of Oregon attribued to the ideal college professor. The results again confirmed the

---

[22]P. L. Crawford and H. L. Bradshaw, "Perception of Characteristics of Effective University Teachers: A Scaling Analysis," Educational and Psychological Measurement, 28 (Winter, 1968), 1083.

[23]Bernadette M. Gadzella, "College Students' Views and Ratings of an Ideal Professor," College and University, 44 (Fall, 1968), 89-96

ideal as one who "(a) stimulates students to think independently rather than to memorize, (b) presents well organized coursework and subject matter, (c) is genuinely enthusiastic about his subject, (d) explains material clearly, and (e) accepts and values student differences in opinion." In contrast, the three least important statements concerning the ideal professor described him as one who "(a) is scholarly as well as an active researcher, (b) has an adequate speaking voice, and, (c) likes and is interested in college age youth as individuals."[24] These findings were supported by work done by Yamamoto and Dizney at Kent State University, noting additionally, no sex differences in the items selected by the students.[25]

In spite of the students' relative lack of concern with research and writing in the ideal teacher images, two sociologists recently tried to determine whether student ratings could be, nonetheless, correlated with actual instructor publications.[26] They developed a

---

[24]Alan F. Quick and Arnold D. Wolfe, "The Ideal Professor," Improving College and University Teaching, 13 (Summer, 1965), 133-134.

[25]Kaoru Yamamoto and H. F. Dizney, " Eight Professors: A Study on College Students' Preferences Among Their Teachers," Journal of Educational Psychology, 57 (June, 1966), 146-150.

[26]William M. Stallings and Sushila Singhal, "Some Observations on the Relationships Between Research Productivity and Student Evaluations of Courses and Teaching," The American Sociologist, 5 (May, 1970), 141-143.

"Research Productivity Index" tabulating faculty publications at two midwestern state universities and correlated it with student ratings. Their findings were inconclusive: "The bulk of the data lend some support to the position that 'publication is not associated with poor teaching.' Conversely, these same data do not offer convincing proof that publication is related to good teaching."[27] In other words, publishing per se was not a function of good or bad teaching, as rated by students, at least.

A study conducted at the University of Toledo proposed specific items to be used in the development of valid ratings schedules.[28] Employing a cross-section sample of students, faculty, and alumni, it "centered on identifying effective teaching behaviors and determining their relative importance."[29] An open-ended technique identified "effective teacher behaviors," and gathered responses from 812 students, 166 faculty, and 665 alumni. A panel categorized the responses into sixty criterion items which were rated on a five-point

---

[27]William M. Stallings and Sushila Singhal, "Some Observations on the Relationships Between Research Productivity and Student Evaluations of Courses and Teaching," The American Sociologist, 5 (May, 1970), 142.

[28]Richard R. Perry, "Criteria of Effective Teaching Behavior in an Institution of Higher Education," Proceedings of the Seventh Annual Forum of the Association for Institutional Research (n. p.: The Association for Institutional Research, 1967), pp. 49-59.

[29]Ibid., p. 51.

scale by another cross-section sample. Significantly, despite the spread of sixty items, compared usually with twenty, the first and last ten items virtually paralleled those of other ideal professor studies involving student samples only. The most important items identified were:[30]

1. Being well prepared for class
2. Establishing student interest in the subject being taught
3. Demonstrating comprehensive knowledge of his subject
4. Using teaching methods which enable students to achieve objectives of the course
5. Constructing tests which search for understanding on the part of the students rather than rote memory ability
6. Being fair and reasonable to students in evaluation procedures
7. Communicating effectively at levels appropriate to the preparedness of students
8. Encouraging intelligent independent thought by students
9. Organizing the course in logical fashion
10. Motivating students to do their best

The least important items of the sixty were:

51. Beginning and ending classes on time
52. Being neatly dressed
53. Being knowledgeable about the community in which he lives
54. Having irritating personal mannerisms
55. Involving himself in appropriate university committees
56. Holding membership in scholarly organizations
57. Being consistently involved in research projects

---

[30]Richard R. Perry, "Criteria of Effective Teaching Behavior in an Institution of Higher Education," Proceedings of the Seventh Annual Forum of the Association for Institutional Research (n. p.: The Association for Institutional Research, 1967), pp. 57-58.

58. Devoting time to student activities on campus
59. Making appearances which assist programs of community organizations
60. Publishing material related to his subject field

Although the Toledo study aimed specifically at validating items for evaluation instruments, other studies attempted to identify those ideal professor items associated with differing discipline areas.

Riley's study also probed those characteristics associated with instructors in different subject areas. The five most frequently chosen attributes for instructors in the arts, sciences and social sciences he found were:[31]

Arts
Knowledge of subject.....54%
Encourages thought..... 47%
Enthusiastic.....46%
Ability to explain clearly..... 42%
Systematic organization
(of subject matter).....32%

Sciences
Ability to explain.....89%
Organization.....78%
Knowledge.....70%
Encourages thought.....17%
Enthusiasm.....16%

Social Sciences
Encourages thought.....70%
Organization.....48%
Tolerance to disagreement..... 45%
Knowledge 42%
Explanation 38%

---

[31]John W. Riley, Jr., Bryce F. Ryan, and Marcia Lifshitz, The Student Looks at His Teacher (New Brunswick, New Jersey: Rutgers University Press, 1950), p. 75.

Even though Riley believed the differences identified with each subject area to be significant, he was as much impressed by "the relatively high degree of similarity in the judgments of different segments of the student body regarding ideal instructional traits. None of the variables studied reveal striking differences in the traits commonly judged most important. Although interesting differences arise, the proportion of students of different types selecting certain ideal traits, there is an over-all homogeneity in popular demand for certain attributes, depending upon the type of course given, but not upon student characteristics insofar as they are measured here."[32] Focusing on the grade differences among students, Turner's study of honors, average, and randomly selected students at Indiana University also found no remarkable difference in their choices of the ideal professor.[33]

Recently Lewis' survey of upper division students at a "large Northeastern university," determined the images students attributed to "three academic occupations (professors of science, social science,

---

[32]John W. Riley, Jr., Bryce F. Ryan, and Marcia Lifshitz, The Student Looks at His Teacher (New Brunswick, New Jersey: Rutgers University Press, 1950), p. 77.

[33]Richard L. Turner, et al., "How Do Student Characteristics Affect Their Evaluations of Instruction," Indiana University School of Education Bulletin, 45 (July, 1969), 48-97.

and humanities) . . . utilizing twenty sets of polar adjectives."[34] Supporting Riley, he reported that among students with dissimilar social or academic backgrounds, "no consistent differences were found [among them] when they were divided by sex, social class, year in college, grade average and extent of extracurricular participation . . . ."[35] Lewis noted, however, "marked differences among those in different majors."[36] Although the results were not uniform, a definite pattern of inversion emerged. For example, students majoring in science and humanities perceived their departmental instructors as being high in congeniality, while each group inverted this image for the other department's instructors.[37] What seemed significant was not that there were differences, such as Riley and Lewis found by subject area, but that the differences tended to be related to instructor behaviors, rather than student characteristics.

Furthermore, these subject area differences could be understood in relation to accepted concepts of role orientations. For within each subject area reported by Riley and Lewis, the ideal

---

[34]Lionel S. Lewis, "Students Images of Professors," Educational Forum, 32 (January, 1968), 186.

[35]Ibid., p. 188.

[36]Ibid.

[37]Ibid., p. 189.

professor traits crossed the full spectrum of role orientations. Thus, despite a certain role dominance, such as instrumental-task ones in the sciences, compared with expressive-integrative ones in the arts, no subject area was identified with one orientation at the exclusion of the other.

Although investigations of the ideal professor served a number of purposes, these studies, along with the ratings ones, were criticized for losing sight of the dynamics of the actual teaching-learning process. Investigators responded by noting their findings produced high degrees of correlation, or perhaps significant factor-loadings for many of the scale items used. Yet critics more often challenged these results on idiosyncratic, not methodological grounds, In disavowing a questionnaire for rating teaching, a dean at Yale University charged:[38]

> As yet we do not know with any degree of positiveness, what combinations of qualities make the successful teacher. There are doubtless various combinations. Of my own best teachers in school and in college one man was rather irritable; one was timid and shy; he limped and walked with a stick; and another was sarcastic with an amusing streak of humor in his sarcasm; another disguised his seriousness with a whimsical manner; another was most intolerant of differing opinions. But we students felt they were all men. They were all competent. Two were great scholars. They were perfectly devoted to their profession.

---

[38]John W. Riley, Jr., Bryce F. Ryan, and Marcia Lifshitz, The Student Looks at His Teacher (New Brunswick, New Jersey: Rutgers University Press, 1950), p. 6.

Similarly, Goheen reflecting upon the elusive nature of teaching, stated that good teaching can never be studied objectively and scientifically because of its complexity and personal nature.[39]

Despite the fact that this type of idiosyncratic criticism was usually dismissed by investigators, these claims were bolstered by researchers supporting student evaluations. Following their survey of the ratings and ideal professor studies, Cohen and Brawer concluded "that the relationship of teacher behavior to student learning is not known and despite decades of research, we have not yet begun to understand those influences."[40] Again, Cohen indicated "that the whole area of teacher evaluation is beclouded with ambiguity and bereft of determinant criteria.[41] According to Shoben, the crux of this problem remained "the remarkable lack not only of any comprehensive theory of teaching, but also of any definitional conception of it that gives unity and meaning . . . "to student evaluations of teaching. He added that in spite of greater sophistication in

---

[39]Robert F. Goheen, "The Teacher in the University," American Scientist, 54 (February, 1966), 221.

[40]Arthur M. Cohen and Florence B. Brawer, Measuring Faculty Performance (Monograph Series; Washington, D. C.: American Association of Junior College, 1969), p. 36.

[41]Arthur M. Cohen and Edgar A. Quimby, "Trends in the Study of Junior Colleges: 1970," Junior College Research Review, 5 (September, 1970), n. p.

statistical techniques and principles of measurement, "there has been little investment of conceptual thought in the teaching process itself."[42]

Reformulating the Evaluations Framework

Reexamining the concerns of the early formulators of ratings research helped clarify this situation, however. In large measure the earlier rationale for student evaluations was closely associated with student expectations. This meant that instructors would learn the expectations of their students from the specific teacher behaviors identified by the end-of-term ratings studies. Once learned, this would contribute to faculty effectiveness. Riley proposed such an evaluations framework initiated by student expectations.[43] With few illusions about the aims and limitations of evaluations, Riley asserted:[44]

> It is not our intention to make any claim in support of the ability of any student body to make an objective and valid analysis of a faculty. That, it must be re-emphasized, is not the point at issue. The real

---

[42]Edward J. Shoben, Jr., "Gimmicks and Concepts in the Assessment of Teaching," Improving College Teaching, ed. Calvin B. T. Lee (Washington, D. C.: American Council on Education, 1967), 292.

[43]John W. Riley, Jr., Bryce F. Ryan, and Marcia Lifshitz, The Student Looks at His Teacher (New Brunswick, New Jersey: Rutgers University Press, 1950).

[44]Ibid., p. 32.

> point, rather, lies in the assumption that student judgments and evaluations, however immature, biased or prejudiced they may be, contribute to the complex of realities in any teaching situation. The professor is dealing with human beings and even in the classroom, where he exercizes a high degree of control and authority he cannot separate these 'beings' from their prejudices and gratifications.

Even though later research established the reliability of students' rating judgment, little concern was expressed for broadening the end-of-course research framework. Riley tried to counter the prevailing trend by proposing a redefinition of the student-instructor relationship:[45]

> Effective teaching can rely on no standardized system of techniques and scholarship; it must take into account the peculiar nature of the student. What the student hears is more important than what the professor says; what the student sees is more important than what the professor does. The student's definition of the professor is as tangible a part of the instructional situation as are the skills and knowledge of the professor. Whatever the goal of instruction, a sound working relationship between student and professor is necessary for the fulfillment of that goal.

While Riley contended that instructors would actually learn student expectations by studying the traits of the ideal professor, he emphasized the interaction of student and teacher, with student expectation inputs playing a continuing role in the dynamics of the classroom. Supporting

---

[45]John W. Riley, Jr., Bryce F. Ryan, and Marcia Lifshitz, The Student Looks at His Teacher (New Brunswick, New Jersey: Rutgers University Press, 1950), p. 33.

this, Riley advanced that:[46]

> An integral complement to the rating scale itself has been the student's expression of the instructional ideals. Without such a yardstick the full significance of the rating on any attribute is lost, for we would not know whether or not the attribute itself held high value for the student. This point takes on added meaning to the individual instructor who studies his own evaluation when we recall that some attributes are undoubtedly 'competitive' with others. Thus if organization is wittingly sacrified for the sake of stimulating insights generated in the classroom, the professor may find it useful to know the significance attached to each of these partially competing attributes. (Italics mine.)

Finally, Riley recognized the explicit consensual relationship between expectations and evaluation:[47]

> If the student is to be given some systematic opportunity to evaluate his professor, he must at the same time be allowed to express his ideal expectations of the professional function. For one without the other becomes only a half statement and may be quite meaningless. There is, for example, one conclusion to be drawn from a student who is critical of a professor who has failed to give what the student most wanted, i. e., stimulation to individual thinking; but quite a different conclusion is demanded if the student expressed little regard for this pedagogical result but rather pinned his hopes on 'learning the subject matter.' (Italics mine.)

Riley's contributions provided an answer to the initial question posed

---

[46] John W. Riley, Jr., Bryce F. Ryan, and Marcia Lifshitz, The Student Looks at His Teacher (New Brunswick, New Jersey: Rutgers University Press, 1950), p. 57.

[47] Ibid., p. 61.

in this chapter: what did evaluations measure? Certainly they measured student role expectations of instructors. It was therefore timely to develop an evaluations perspective that went beyond end-of-course instructor ratings. A more inclusive and explicit evaluative framework for college instruction proved necessary, a framework that included student expectations.

CHAPTER II

A FRAMEWORK FOR STUDENT ROLE EXPECTATIONS OF INSTRUCTORS

Despite Riley's proposals, few evaluations researchers examined the problem of how students entering a particular course orient themselves to the instructor. Moreover, few studies investigated student definitions of their role-oriented tasks, especially in association with the classroom, curriculum and faculty; in a word, an "evaluations ambience."[1] Yet no investigator challenged the assertion, that "the students' orientation, their expectations and role definitions are likely to provide considerable insight into how the students will adapt to the demands made upon them."[2]

Even with the increased investigations related to student expectations of the overall college environment, the absence of

---

[1]W. J. McKeachie, "Research on Teaching at the College and University Level," Handbook of Research on Teaching, ed. N. L. Gage (Chicago: Rand McNally and Company, 1963), 1118-1172. Cf. D. Bob Gowin and Donald E. Payne, "Evaluating Instruction: Cross Perceptions of College Students and Teachers," School Review, 70 (Spring, 1962), 217-219.

[2]F. M. Katz and Cecily N. Katz, "Great Expectations," Universities Quarterly, 23 (Autumn, 1969), 420.

research concerned with student expectations of instruction persists. Researchers seemed to take for granted that students' total expectations of courses and instructors contributed to the successes and failures in their college experience, without apparent evidence.[3] Although such college environment studies suggested "that people work more effectively in situations that conform to their preferences,"[4] there still remained inadequate evidence explaining expectations held for specific courses and instruction. College environment studies indicated, nevertheless, that if a student's expectations were realistic, the process of adaption to the college environment became relatively smooth, and was likely to be reflected in successful fulfillment.[5] This point served additionally as a reason to examine expectations related to courses and instructors. Since this study assumed, however, that the student dimension related to classroom expectations was a broad one, it concerned itself only with student role expectations of college teachers in relation to student evaluations, in a particular course of study.

---

[3]F. M. Katz and Cecily N. Katz, "Great Expectations," Universities Quarterly, 23 (Autumn, 1969), 420. See also Lawrence A. Pervin, "Reality and Non-reality in Student Expectations of College," Journal of Psychology, 64 (September, 1966), 41-48.

[4]Ibid., p. 420.

[5]Kenneth A. Feldman and Theodore M. Newcomb, The Impact of College on Students (San Francisco: Jossey-Bass, Inc., 1969), I, 81-82. See K. A. Feldman, "Studying the Impact of Colleges on Students," Sociology of Education, 42 (Summer, 1969), 207-237.

To be sure, the present study could only be a first approximation of student expectations of specific courses and instructors. Accordingly, it examined the relationship between student expectations in certain classes at the start of a semester compared with their subsequent ratings toward the end of that same semester. This procedure contrasted with the typical research design of student evaluations.[6] Invariably these investigations sought student responses on instruments designed to rate the instructor upon--or near--the completion of a course. From the collected data, researchers drew their conclusions concerning the influence of intervening variables on their evaluations. Operationally, this meant that a student's initial expectations were to be inferred from an end-of-term rating. From the point of view developed in this paper, end-of-course ratings should be examined as an integral part of an evaluative process that includes measuring initial student expectations.

The Classroom as a Consensual Experience

Newcomb's formulation of "consensual expectations" was most useful in making this evaluative process more inclusive.[7] Although

---

[6]H. H. Remmers, "On Students' Perceptions of Teachers' Effectiveness," The Appraisal of Teaching In Large Universities, ed. W. J. McKeachie (Ann Arbor: University of Michigan Press, 1959), 17-20.

[7]Theodore M. Newcomb, "Student Peer-Group Influence," The American College, ed. Nevitt Sanford (New York: John Wiley and Sons, 1962), 469-488.

Newcomb concerned himself with shared peer group influences in relation to learning outside of the classroom, his formulation served as a bridge connecting student expectations and teacher objectives. Borrowing this model of shared relationship for the teaching-learning framework proved to be a working model for classroom research. (If developed as a research model, it might become an integral factor in the process of student-college adaptation and outcomes.) In comparison with this open-ended framework, Rudolph expressed deep concern for the traditional exclusion of students from the teaching-learning process both in practice and research: "And what is most distressing of all is how often in our history students have had to tell us of their presence--of their needs as young human beings discovering the limits of their individual destinies."[8] It seemed appropriate, therefore, to develop an evaluations research model focusing primarily on student inputs. Unlike research models characterizing student inputs as achievement or performance levels, intelligence quotients, or socio-economic status, the consensual expectations idea suggested that initial expectations added to a more inclusive understanding of the student's domain, especially in relation to their evaluations of college courses.

One problem faced by employing the consensual expectations

---

[8]Frederick Rudolph, "Neglect of Students as a Historical Tradition," The College and the Students, eds. L. E. Dennis and J. F. Kauffman (Washington, D. C.: American Council on Education, 1966), 58.

approach as an evaluations model was the limited empirical findings on shared student-instructor interaction in the classroom. Surely no researcher--or practitioner--overtly denied the centrality of this relationship, actual studies of student expectations of college had usually considered non-instructional characteristics, such as dormitory regulations or social and extracurricular activities, treating the academic variables in more panoramic terms. Even the widely used College Characteristics Index developed by Stern and Pace contained faculty-related items of so general a nature as to be of little value in discussing specific behaviors related to instruction. The Index used such items as: "Some of the professors treat questions in class as if the students were criticizing them personally," or "Many faculty members seem moody and hard to figure out."[9] These items certainly related to the perception of an overall environment, but added little understanding of specific student expectations of instructors. Used to demonstrate increasing student disaffection and alienation from college,[10] these studies did not measure whether failure to meet student instructional expectations was connected to instructors and courses as such, and not with generalized institutional dynamics. Instead, consensual

---

[9]George C. Stern, People in Context: Measuring Person--Environment Congruence in Education and Industry (New York: John Wiley and Sons, 1970), p. 343.

[10]George C. Stern, "Myth and Reality in the American College," AAUP Bulletin, 52 (December, 1966), 408-414.

expectations suggested that student instructional evaluations could be more appropriately understood when treated as a function of their expectations.

Correspondingly, the research aim of consensual expectations suggested expanding current notions of student evaluations of instructors.[11] Here it followed recent models associated with innovations in undergraduate education, specifically those models emphasizing teacher-learner interaction.[12] Usually organizational theory, most often the process or systems models, supplied the conceptual framework for various innovative strategies.[13] By contrast, previous evaluations studies presumed the equilibrium model of classroom organization. This unidimensional approach operated with the least possible disturbance: the instructor taught and the students learned. Consequently, end-of-course ratings were not conceived as reciprocal or mutually derived task fulfillments, but measured teacher behavior

---

[11]Joseph Axelrod, Model Building for Undergraduate Colleges, U. S. Office of Education Contract No. OEC 6-10-106 (Washington, D. C.: Educational Resources Information Center, 1969), pp. 1-32.

[12]Douglas H. Heath, Growing Up in College: Liberal Education and Maturity (San Francisco: Jossey-Bass, Inc., 1968), pp. 20-48.

[13]Joseph Axelrod, Model Building for Undergraduate Colleges, U. S. Office of Education Contract No. OEC 6-10-106 (Washington, D. C.: Educational Resource Information Center, 1969), pp. 1-32 and passim.

exclusively. The process model or complex-adaptive system, thriving on "disturbances and variety" in the environment, offered an evolving framework with potential for high order student and teacher interchange.[14] If the process model contributed this concept of interchange to consensual expectations, the systems model proposed the concept of feedback. Essentially this meant that student and teacher monitor and reassess each other's expectations and objectives regularly. Taken together, the process and systems models permitted student involvement in the articulation of instructional objectives, in addition to producing instructor awareness of student dispositions, with some provision for feedback. Typical end-of-term ratings instruments might be included, too. But not without student and teacher becoming mutually responsible for the fulfillment and assessment of each other's stated and mediated expectations and objectives. In these terms, then, the structure of consensual expectations would be determined largely by interchange and feedback. The evidence demonstrated that this task should be realizable for both large and small institutions. For, as Eddy's findings in 1959 indicated, which Feldman and Newcomb restated in 1969: "The extent to which the student comes into direct contact with the professor depends more upon the attitudes and efforts

---

[14]Walter Buckley, Sociology and Modern Systems Theory (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1967), p. 40.

of the institution than upon its size."[15]

Understandably, there would be some lag between the growth of these role-sharing typologies in research and their acceptance in practice. In fact, in anticipating possible colleague criticism, investigators of student evaluations often assumed a defensive research posture.[16] In 1961, and again in 1967, Gustad remarked that faculty criticism persisted even after end-of-term ratings appeared on more college campuses.[17] In their recent study Cohen and Brawer reported that "a somewhat cynical opinion among some teachers that very little value can be placed on student . . . " judgment still prevailed.[18]

Proponents of student ratings contended that faculty criticism stemmed largely from the belief that students lacked maturity to formulate sound evaluative judgments. Two major criticisms of

---

[15]Kenneth A. Feldman and Theodore M. Newcomb, The Impact of College on Students (San Francisco: Jossey-Bass, Inc., 1969), I, 580.

[16]Charles E. Gray, "The Teaching Model and Evaluation of Teaching Performance," The Journal of Higher Education, 40 (November, 1969), 636-642.

[17]John W. Gustad, "Evaluation of Teaching Performance: Issues and Possibilities," Improving College Teaching, ed. Calvin B. T. Lee (Washington, D. C.: American Council on Education, 1967), 276-277.

[18]Arthur M. Cohen and Florence B. Brawer, Measuring Faculty Performance (Monograph Series; Washington, D. C.: American Association of Junior Colleges, 1969), p. 11.

of student evaluations summarized by McKeachie were:[19]

1. Students really can't evaluate a teacher until they've left college and gotten some perspective on what was really valuable to them.
2. Students rate teachers on their personality--not on how much they've learned.

Though critical, Megaw offered an explanation for faculty resistance: "Is it not unaccountable that so many people so openly committed to enlightenment should for so many years put up with such ignorance about a procedure [evaluation] so close to the heart of their common endeavor?"[20] He disclosed two "main faculty motives . . . fear and laziness," explaining "the nightmarish fear of being declared incompetent, or at least shamefully inexpert . . . . " Regarding laziness as less intense than fear, but "more endemic," he continued, "not a general, undifferentiated laziness, however--most teachers put in a long working day--but a special laziness of the experimental spirit: reluctance, in short, to consider new patterns of overwork." In defending "the honorable faculty motive for resisting" evaluation, he added, "[their] loyalty to that which goes on in the classroom is vastly more complex than any definition of it, and that its chief

---

[19]W. J. McKeachie, "Student Ratings of Faculty," AAUP Bulletin, 55 (December, 1969), 439-444.

[20]Neill Megaw, "The Dynamics of Evaluation," Improving College Teaching, ed. Calvin B. T. Lee (Washington, D. C.: American Council on Education, 1967), 282.

values reside in this complexity."[21] If Megaw's explanation, as noted by others,[22] that currently understood end-of-term ratings erode classroom complexity, they deserved the criticisms received.

By contrast, the consensual expectations approach would require measurement of instructor evaluations as a function of initial student expectation inputs. If the research being proposed by consensual expectations demonstrated that students were reliable classroom participants, then certain faculty objections to this evaluative process might be dispelled. Furthermore, consensual expectations would certainly add, not detract, from the classroom complexity cited by Megaw as a faculty objection to evaluations.

Consensual Expectations and Group Dynamics Theory

The theories of group dynamics provided the conceptual framework for this research on student expectations. Clearly, if any theory were to make a contribution, it must relate expectations to the role relationships in the college classroom. Despite its concern with an

---

[21] Neill Megaw, "The Dynamics of Evaluation," Improving College Teaching, ed. Calvin B. T. Lee (Washington, D. C.: American Council on Education, 1967), 283.

[22] Kenneth E. Eble, The Recognition and Evaluation of Teaching, Report of the Project to Improve College Teaching, Salt Lake City, Utah, November, 1970 (Salt Lake City, Utah: Project to Improve College Teaching, 1970), pp. 18-19.

industrial setting, Stogdill's work on group achievement provided the appropriate framework.[23] Like Newcomb's consensual expectations framework, Stogdill's work has been identified in the literature of the field as systems and open-ended theory. For Stogdill group performance and achievement were determined by initial member inputs, that is, expectations.[24] Hence, his expectations formulation served as a working model for the determinants of student expectations.

Influenced by earlier sociological theorists such as Mead,[25] Stogdill's definition of expectation could be traced to the work of the learning theorists, Tolman and Kelly.[26] Stogdill defined expectation as "a readiness for reinforcement" which "is a function of drive, the estimated probability of occurrence of possible outcome, and the estimated desirability of the outcome."[27] By drive he meant "the level of tension reactivity exhibited by an organism," which would

---

[23]Ralph M. Stogdill, Individual Behavior and Group Achievement (New York: Oxford University Press, 1959).

[24]Dorwin Cartwright and Alvin Zander, "Introduction," Group Dynamics: Research and Theory, eds. Dorwin Cartwright and Alvin Zander (3d ed.; New York: Harper and Row, Publishers, 1968), 26.

[25]Stogdill, op. cit., p. 5.

[26]Ibid., pp. 60-62.

[27]Ibid., p. 62.

become activated into readiness for reinforcement defined as "the extent to which an individual is prepared or unprepared to experience, or reconciled to the prospect of experiencing, a possible outcome."[28] Actual outcomes would be mediated by an individual's estimates of occurrence which "refers to the individual's prediction, judgment, or guess relative to the likelihood that a given event will occur, "as well as by the estimated desirability which he defined as "an individual's judgment relative to the satisfyingness of, need for, demand for appropriateness of, or unpleasantness of, a possible outcome."[29] In the classroom these two estimates determined student expectations associated with projected outcomes of teacher behaviors. Aware that some question might be raised regarding the positive or negative valence of the estimates, Stogdill added that the "estimate of probability and estimates of desirability are not opposite ends of the same continuum . . . ."[30] Instead, these estimates "interact to determine the level of expectation."[31] And most relevant to the classroom, "this interaction is formulated in terms of what is uncertain in the future as

---

[28]Ralph M. Stogdill, Individual Behavior and Group Achievement (New York: Oxford University Press, 1959), p. 62.

[29]Ibid.

[30]Ibid., p. 128.

[31]Ibid.

well as in relation to what has been learned in the past."[32]

If this concept of expectations was understood as an interaction of probable and desired estimates, the particular behavior outcomes were dependent upon the classroom role structure. Accordingly, Stogdill first distinguished between the place of the formal structure of groups, making it possible to determine the different expectations associated with individual instructors. Stogdill then discriminated between expectations "attached to position rather than the occupant of a position." Since the classroom's formal structure prescribed the professor's status and function which "are defined for a position rather than for any given occupant of the position,"[33] Stogdill emphasized the relevance of expectations related to role structure. Role structure, in short, was the more inclusive concept. By distinguishing between formal and role structure, Stogdill's approach contributed additionally to evaluations research. For if the ideal professor studies identified classes of behavior states associated with the formal instructional expectations, then the study of their interplay in actual classroom settings was one to which this study addressed itself.

Furthermore, in keeping with previously developed organization theory, Stogdill introduced the concept of responsibility and authority

---

[32]Ralph M. Stogdill, Individual Behavior and Group Achievement (New York: Oxford University Press, 1959), p. 123.

[33]Ibid., p. 123.

to differentiate two levels of role structure expectations. Responsibility, he defined as the set of performances that a given occupant was expected to exhibit "by virtue of the operational demands made upon his position," contrasted with authority representing the degree of freedom that the occupant of a role could be expected to exercise as an individual.[34] In effect, both the probable and desirable estimate of behaviors defined by the classroom's role structure would produce "a set of expectations" which "imply, not a psychic entity of any sort but a mutually confirmed readiness for reinforcement . . . ."[35]

Student expectations as proposed by Stogdill were used in this paper as a role-related set of probable and desirable estimates of instructor behaviors for which students exhibited a readiness for reinforcement. This formulation permitted classroom role expectations--and subsequent evaluations--to be understood as both a function of the normative aspects of a classroom coupled with individual instructor characteristics. The interaction of student estimates of desirable and probable instructional behavior outcomes, in other words, not only determined evaluations, as the ideal professor studies suggested, but also activated student expectations. Thus, the concept of expectations employed here included the readiness for projected behavior

---

[34]Ralph M. Stogdill, Individual Behavior and Group Achievement (New York: Oxford University Press, 1959), p. 129.

[35]Ibid., p. 128.

outcomes.[36]

It was obvious that the model adopted by the end-of-course ratings studies limited the understanding of expectations in relation to evaluations. By excluding estimates of probability and desirability, these studies failed to recognize that student expectations precede their evaluations. According to the definition of expectations employed here, an affirmative or negative rating in an instructional behavior category could have exceptional meaning if initial expectation inputs were taken into account. In the ratings studies, however, there seemed to be no way to ascertain whether students expected a teacher to be friendly, even though he in fact turned out to be quite friendly. The students then rated him affirmatively. The student or class could be, nonetheless, expecting to achieve more organized knowledge about the subject than the instructor gave for which students rated him negatively. In the ratings studies the tendency would be to equate the two results, perhaps one equalizing or cancelling the other. Extrapolations of student expectations made from their evaluations could be, in sum, of limited value.

Building upon this perspective, Stogdill's formulation as open-ended group theory contributed to an understanding of the consensual expectations point of view. Here Stogdill's relationship with evaluations

---

[36]Joseph Berger, Thomas L. Conner, and William L. McKeown, "Evaluation and the Formation and Maintenance of Performance Expectations," Human Relations, 22 (December, 1969), 481-502.

should be noted. For if the systems-process models suggested that evaluations are a function of interchange and feedback based on initial inputs, then Stogdill's notion of role structure offered an explanation of the organization of expectation inputs. Analogously, if the systems approach explained the structural and interactive dimensions of consensual expectations, then Stogdill defined the situation that determined expectations.

A recent study by Phillips developed procedures resembling those suggested by consensual expectations approach.[37] As such, it represented the only one attacking the problem of student evaluation as a function of desired instructional outcomes compared with actual ratings. Phillips noted that:[38]

> The outcomes of teaching would seem to be complexly determined by at least four factors: the characteristics of the teacher, the students, the subject matter, and the class as a group. And, what is more important, it appears to be the interaction of these factors which partly produces differences in outcomes in teaching.

Adding that since "the importance of a particular criterion of teaching effectiveness is likely to vary from student to student, and from class to class," Phillips maintained the end-of-term ratings studies, "may

---

[37]Beeman N. Phillips, "The 'Individual' and the 'Classroom Group' as Frames of Reference in Determining Teacher Effectiveness," Journal of Educational Research, 58 (November, 1964), 128-131.

[38]Ibid., p. 128.

often provide an overgeneralized estimate of a teacher's effectiveness."[39] To change the direction of this research, he asked students in eight sections of an educational psychology course toward the end of the term to complete two instruments, one to determine "what happened in this course," and the other to determine which experiences students "considered to be most important in the courses generally."[40] For purposes of analysis, Phillips clustered the instrument items according to studies done at the University of Michigan in which four classroom factors were derived using factor analytic techniques with a relatively large sample of students.[41] These factors were: amount of structure provided; amount of information given about tests; amount of achievement motivation aroused; and, amount of personal warmth in class. By clustering items in this way, Phillips noted that a more realistic picture of a student's expectations would be achieved, as opposed to a more generalized picture usually given by an item by item analysis. He reported significant results at the .01 level in all four categories: students expected more structure and achievement motivation arousal which their experience did not reinforce, and students expected less information about tests and

---

[39]Beeman N. Phillips, "The 'Individual' and the 'Classroom Group' as Frames of Reference in Determining Teacher Effectiveness," Journal of Educational Research, 58 (November, 1964), p. 128.

[40]Ibid., p. 129.

[41]Ibid., p. 128.

personal warmth in class which was overly reinforced by the instructors.[42] Unfortunately, his design was not longitudinal because the two tests were taken successively, and limited because he asked students to use an ideal course framework for the expectations instrument instead of the actual course. Phillips, nonetheless, pointed in the direction of further explanation in expectation-evaluation studies. In anticipating the consensual expectations model, Phillips saw the need for additional research of a longitudinal type.

Phillips' study dealt with the interactive qualities of expectations and evaluations, another one by Twa concerned itself with role conditions that determine expectations.[43] It examined the expectations of community college students in relation to different instructional roles, namely, transfer instructors, occupational instructors, and adult education instructors. Twa contributed a persuasive argument for utilizing the consensual expectations framework:[44]

---

[42]Beeman N. Phillips, "The 'Individual' and the 'Classroom Group' as Frames of Reference in Determining Teacher Effectiveness," Journal of Educational Research, 58 (November, 1964), p. 131.

[43]R. James Twa, Student and Instructor Expectations of Community College Instructors (Eugene, Oregon: Oregon School Study Council, 1970).

[44]Ibid., p. 28.

> Persons have expectations of others with whom they must interact to attain their goals. Conformity to these expectations normally result in rewards, and non-conformity normally results in negative sanctions being applied to the person. However, in the student-instructor situation a complication arises because the students may jeopardize their chances of goal attainment if they (as subordinates) apply negative sanctions to instructors (superiors). Thus, because of their inability to take action, the student's frustration is heightened.

Twa concluded that the teaching-learning relationship is a consensual one: "Each is dependent upon the other to fulfill successfully the objectives for his respective mission."[45] The findings of his study disclosed that students place great importance on the personal relationships they expect of instructors.[46]

Relating Classroom and College Expectations

Despite its concern with institutional and environmental variables, Stern's work added another dimension to the concept of expectations.[47] Stern developed a College Characteristics Index to determine student expectations of the demands or pressures of a college's

---

[45]R. James Twa, Student and Instructor Expectations of Community College Instructors (Eugene, Oregon: Oregon School Study Council, 1970), p. 5.

[46]Ibid., p. 26.

[47]George C. Stern, People in Context: Measuring Person--Environment Congruence in Education and Industry (New York: John Wiley and Sons, 1970).

environment, called press.[48] The Index also permitted Stern to determine aspects of individual personality, called needs.[49] Accordingly, the needs scale attempted to identify "the goals or purposes that an interaction serves for the individual. In this sense a listing of needs is essentially a taxonomy of the objectives that individuals characteristically strive to achieve for themselves."[50] The concept of environment press, as a corollary of needs, was defined by Stern as "the phenomenological world of the individual, the unique and inevitably private view each person has of the events in which he takes part."[51] In order to differentiate between an individual's press expectations, and those presumably posited in the institution, Stern made a distinction between "explicit objectives, representing the stated purposes for which given institutional events are organized, and the implicit objectives, which are in fact served by institutional events."[52] Thus the taxonomy of a student's expectations, which the Index measured for given institutional elements,

---

[48]George C. Stern, People in Context: Measuring Person-Environment Congruence in Education and Industry (New York: John Wiley and Sons, 1970), p. 7.

[49]Ibid.

[50]Ibid., p. 6.

[51]Ibid.

[52]Ibid., p. 7.

such as faculty, could be different from a "detached observer."[53] That is, a student's expectations of "the situational climate, the permissible roles and relationships, the sanctions and so on, "could differ from the detached observer, presumably the investigator.[54] To apply this typology in practice, Stern suggested that the congruence or dissonance of student expectations could be compared with the reality of an institution's matrix.

Regarding the curricular and academic expectations of college freshmen, some additional questions the Index raised related to faculty should be noted: were students treated formally or informally by faculty? Were faculty demands upon students heavy or light? Did the general teaching procedure emphasize lectures or free discussions? Stern's use of the Index on a large sample of freshmen at Beloit College, Cazenovia College, St. Louis University, and Syracuse University, seemed most relevant here.[55] He advanced the proposition that a "Freshman Myth" existed in American colleges, for they "share stereotyped expectations of college life that combine some of the elite liberal arts college with the community spirit, efficiency and social orderliness of

---

[53]George C. Stern, People in Context: Measuring Person--Environment Congruence in Education and Industry (New York: John Wiley and Sons, 1970), p. 7.

[54]Ibid.

[55]Ibid., p. 92.

church-related schools."[56] More emphatically, he declared: "University-bound high school seniors evidently share a highly idealized image of college life representative of no actual institution at all."[57] Moreover, despite the fact that the four colleges in his sample were quite different, "the expectations of the four groups of freshmen follow a substantially similar pattern."[58] These freshmen "look forward to high levels of activities relevant to both academic and nonacademic press . . . ."[59] Yet, "this does not correspond to the actual characteristics of these schools at all. It is evident that the incoming freshman expected something rather different from his upper division colleagues . . . . As an entering freshman, he came expecting to learn, as a senior he has learned not to expect so much."[60]

In order to consider these findings a distinction had to be made between the two differing dimensions of expectations being treated. The expectations dimension examined by Stern, associated with the

---

[56]George C. Stern, People in Context: Measuring Person--Environment Congruence in Education and Industry (New York: John Wiley and Sons, 1970), p. 173.

[57]Ibid.

[58]Ibid., p. 92.

[59]Ibid.

[60]Ibid.

over-all press of an institution, should be called situational expectations. By comparison, the dimension of expectations put forward by Stogdill and represented by consensual expectations, which were linked with group role structures, should be called contextual expectations. Even if Stern differentiated between explicit and implicit group objectives, resembling Stogdill's formal and role structure expectations, these objectives went beyond reference to specific classroom role contexts.

This distinction made a comparison between the two concepts possible: did freshmen contextual expectations differ significantly from other classmen? In turn, were the subsequent freshmen evaluations significantly different from their expectations lending support to the incipent alienation thesis Stern projected as part of a general freshmen disenchantment? Moreover, did the freshmen exhibit the remarkably similar contextual expectations of their instructors as Stern's sample exhibited for the over-all situational ones? Answers to these questions presented by this study could invite further research concerning contextual and situational expectations in varying aspects of college culture.

CHAPTER III

DESIGN, RESEARCH PROCEDURES AND HYPOTHESES

As a study of student contextual expectations of college instructors, the design was determined by the natural setting of the classroom. For college classroom research Gage distinguished three types of appropriate study designs.[1] First, descriptive studies depicting aspects of the teaching-learning process. Second, experimental research manipulating certain aspects of the classroom situation--the independent variable--in seeking measurable changes in students. And conversely, experimental research relating changes in student--the dependent variable --to some measure of the teacher's behavior or classroom experience.[2] In adopting the latter method, this study treated expectations as the independent variable, and evaluations and certain other variables to be noted later, as the dependent ones.

Study Design and Procedures

The study also used a one-semester longitudinal design. At

---

[1]N. L. Gage, "The Appraisal of College Teaching: An Analysis of End and Means," Journal of Higher Education, 32 (January, 1961), 17-22

[2]Ibid., p. 21

the beginning of a semester a sample of college students reported their expectations, and after an interval, their evaluations were determined. This method avoided the possible instrument contamination noted regarding Phillips' study of testing and posttesting on the same day.[3] Furthermore, since the present study treated evaluations as a function of expectations, the posttest could have been given at any point during the semester.

Arrangements were made in the Spring of 1970 in three colleges and with eight instructors in whose introductory social science classes this study was conducted. Selection of the instructors was based on two considerations. One, the subject matter in each course was relatively equivalent, and, two, the instructors employed discussion methods predominantly. The author determined this point by a prior classroom visit and discussion with the participating instructors. Moreover, an introductory course was chosen because it enrolled a high percentage of freshmen, making any finding regarding their contextual expectations more representative. Divulging only that the study aimed to learn more about students taking social science courses, the author requested the instructors not to discuss the study or the instruments, which remained unseen, with either students or colleagues. Having also agreed not to reveal to their students that they would participate

---

[3]Beeman N. Phillips, "The 'Individual' and the 'Classroom Group' as Frames of Reference in Determining Teacher Effectiveness," Journal of Educational Research, 58 (November, 1964), 128.

in this study, each instructor introduced the author on the second day of class, and then left the room; the expectations instrument was then administered. Students were not told that the instrument would be retaken with modifications for evaluations purposes following ten weeks of classes. Insofar as can be determined, the instructors cooperated fully by not discussing the study with students. In two classes with absentees on the posttest, the instructors asked the author to return on the next class day so the students could complete the schedule.

This longitudinal design was coupled with an interinstitutional sample. Although most evaluations research was based on single-institutions, Astin recently advanced that the multi-institutional approach acted as a modified control, and should therefore be encouraged.[4] The three colleges were chosen primarily because of their differences in character and clientele. Chosen because they shared the explicit institutional goal of stressing the teaching function, they were all, to be sure, located conveniently within a fifty-mile radius. Alpha College, a seventy-year old two-year women's residential college with a somewhat selective enrollment close to 400, offering both liberal arts transfer and terminal programs in the fine, applied and performing arts, and early childhood education, had four of its classes surveyed. About 25 students were in each class. Beta College, by contrast, a comprehensive community college about a dozen years old with a full-time enrollment of

---

[4]Alexander W. Astin, "The Methodology of Research on College Impact," Part I, Sociology of Education, 43 (Summer, 1970), 251

2,500, having an open-admissions policy had two classes surveyed. Both had about 25 students. In comparison with these colleges, Gamma College, a one-hundred year old four-year women's residential college, recently turned to co-educational, enrolling 200 males out of 1,600 students, with some strong liberal arts programs, and a highly selective admissions policy, had two classes surveyed. One enrolled 25 students, the other 40. Owing to the full cooperation of the instructors, all students completed both the expectations and evaluations instruments, resulting in a total student population of 209.

Developing an Instrument

The instrument employed in the study developed principally from two different analytical techniques. These techniques were used to transform the categories developed in ratings and ideal professor studies into specific instrument items. Perhaps the most widely selected items were found in the University of Michigan studies which applied factor-analytic techniques to items gathered from previous study instruments.[5] From a pool of 145 items, the Michigan studies derived 34 items through rotated factor analysis.[6] It may be recalled that Phillips

---

[5]Robert L. Isaacson, Wilbert J. McKeachie, and John E. Milholland, "Correlation of Teacher Personality Variables and Student Ratings," Journal of Educational Psychology, 54 (April, 1963), 110-117, and R. L. Isaacson, et al., "Dimensions of Student Evaluations of Teaching," Journal of Educational Psychology, 55 (December, 1964), 344-351.

[6]R. L. Isaacson, et al., op. cit., p. 345.

utilized four instructional categories which he derived from these 34 items.[7]

The other attempt to distinguish as well as to cluster items within instructional categories employed the Critical Incident technique developed in a University of Houston study.[8] Because its approach actually identified items related to contextual expectations of college instructors, rather than evaluations per se as the Michigan study did, its usefulness to the present study was obvious.

Based on previous work by Flanagan, Owen defined the Critical Incident technique as "a systematic method of observing and analyzing human behavior."[9] Essentially phenomenological in approach, Flanagan defined an incident as "any observable human activity that is sufficiently complete in itself to permit inferences and predicitions to be made about the person performing the act."[10] To account for situational exceptions, Flanagan added: "To be critical, an incident must occur in a situation where the purpose or intent of the act seems fairly clear to the observer and where its consequences are sufficiently definite

---

[7]Beeman N. Phillips, "The 'Individual' and the 'Classroom Group' as Frames of Reference in Determining Teacher Effectiveness," Journal of Educational Research, 58 (November, 1964), 128-131.,

[8]Paul H. Owen, "Some Dimensions of College Teaching: An Exploratory Study Using Critical Incidents and Factor Analysis of Student Ratings" (unpublished Doctoral dissertation, University of Houston, 1967).

[9]Ibid., p. 19.

[10]Owen, op. cit., p. 21.

to leave little doubt concerning its effect."[11] Used extensively in studies of elementary and secondary school teachers, notably in Ryans' nationwide survey on teacher characteristics,[12] Owen employed structured and unstructured procedures to obtain the initial inventory of critical incidents that students associated with college instructors' behaviors. From the resulting list of behaviors, Owen developed an instrument, which he subsequently administered along with the Michigan instrument. His findings demonstrated a high degree of intercorrelations between the items obtained through the critical incident and Michigan factor analysis techniques. Likewise, the six categories he turned into clusters from the Critical Incident techniques were "substantially verified . . . " with the Michigan categories.[13]

Fortunately, these findings offered the first empirical evidence that students employed similar behavioral frameworks for their expectations and evaluations of instructors. Consequently, the current study clustered six categories identified by the Houston Critical Incident method for purposes of analysis, in addition to selecting instrument items from it and the Michigan factor-analysis technique studies. To

---

[11]Paul H. Owen, "Some Dimensions of College Teaching: An Exploratory Study Using Critical Incidents and Factor Analysis of Student Ratings" (unpublished Doctoral dissertation, University of Houston, 1967), p. 21.

[12]David G. Ryans, Characteristics of Teachers (Washington, D. C.: American Council on Education, 1960), pp. 79-83.

[13]Owen, op. cit., p. 140.

clarify the major classes of instructor behaviors identified through the Critical Incident technique,[14] a description of the six categories follows:

I. Presentation of Material-Content, Structure and Scope

Behaviors related primarily to the structuring of the content. These include the organization, planning, selection and preparation of content; use of supplementary references and illustrations; use of practical examples including personal experiences; thoroughness of explanations and level of difficulty of presentation; apparent knowledge of subject.

II. Presentation of Material-Student Participation

Behaviors related primarily to student involvement in presentation of material. These include instructors' relative emphasis on lecture and/or class participation, student involvement in organizing and presenting material and assignments to students which specifically relate to the presentation of material.

III. Presentation of Material-Instructor's Style

Behaviors related primarily to the instructor's individual style and choice of techniques of presentation. These include level of enthusiasm for the subject and its presentation; animation; use of humor; speech characteristics; rate of presentation; use of visual aids; individual presentation techniques and traits.

IV. Teacher-Student Rapport and Class Interaction

Behaviors related more to affective components of instructor and student interaction than to subject-oriented student participation. These include the instructor's approach to formality of class; social distance between teacher and student; permissive versus authoritarian style, personal interest in and involvement with students and their problems; personality characteristics to which students

---

[14]Paul H. Owen, "Some Dimensions of College Teaching: An Exploratory Study Using Critical Incidents and Factory Analysis of Student Ratings" (unpublished Doctoral dissertation, University of Houston, 1967), p. 47.

react; control and discipline in class.

V. Evaluation of Students

Behaviors related primarily to the processes surrounding the appraisal of students' progress. These include the adequacy of defining test requirements; practices and procedures in grading; type, frequency and content of tests.

VI. Requirements of Students

Behaviors related primarily to what is expected of students but excluding those having to do with assignments to present material. These include adequacy of defining course requirements; responsibilities given to, demands made of outside assignments and expectations of students.

Before noting the actual instrument items and discussing scaling technique, the validity issue related to expectations-evaluations instrumentation should be discussed.

Even though the Michigan and Houston studies validated certain items, researchers implicitly assumed that virtually any item used on an evaluations instrument was self-validating. Often studies made little attempt to cite or report previously developed instructional categories or items, apparently assuming that the instructional areas were adequately covered.[15] Since most researchers tended to look favorably upon student evaluations, the reluctance to explore instructional categories more thoroughly could be explained in terms

---

[15]See Richard R. Renner, "A Successful Rating Scale," Improving College and University Teaching, 15 (Winter, 1967), 12-14, and David A. Strand, "A Rationale and Instrument for Student Evaluation of Classroom Teaching," Journal of National Association of Women Deans and Counselors, 30 (February, 1966), 36-39.

offered by a student advocate of instructor ratings, that "there are no bad student course and teacher evaluations."[16] More explicit, Remmers asserted that, "by definition, if one is concerned only with measuring the perception the students have of instructors, validity equals reliability."[17] In the handbook accompanying the Purdue Rating Scale for Instructors with which Remmers is associated, the instructions noted, "there is, of course, no easily available outside criterion for this. Since the students' attitudes and perceptions are at issue, validity by definition can be equated with reliability."[18] This would certainly be the case if the high degree of intercorrelation of the Michigan and Houston items were acknowledged. Indeed, when separate classes at Alpha College retook the current study's instruments, one after a 10 day and the other a 14 day interval, they produced on its seven-point scale reliability coefficients of .89 and .87 respectively.

---

[16]Philip R. Werdell, Course and Teacher Evaluation (2d ed.; Washington, D. C.: National Student Association, 1966), p. 5, underlined in the original.

[17]H. H. Remmers, "On Students' Perceptions of Teachers' Effectiveness," The Appraisal of Teaching in Large Universities, ed. W. J. McKeachie (Ann Arbor: University of Michigan Press, 1959), 21.

[18]H. H. Remmers, Manual of Instructions for the Purdue Rating Scale for Instruction (Rev. ed.; Lafayette Indiana: Purdue University Book Store, 1960), p. 2.

Scaling Technique Employed

Developing an appropriate scaling technique for this instrument proved difficult. There was little consistency in scaling techniques for evaluation instruments. Even when studies utilized items from earlier investigations, no uniform pattern of scaling procedures emerged. Most investigators applied some form of the forced-choice technique with scales ranging from five to ten points; five being the most common. Yet, if one study adopted a nine-point sentence completion technique with the opposite poles being some variation of, "I found the textbook in this course to be: 1)Very dull" to "9) Very interesting,"[19] another used adjectives such as "probing" or "sympathetic" on an open-ended seven point scale.[20] Some studies mixed quantitative and qualitative measures on one instrument. As an example, using a closed five-point scale for the item, "Class time is: usually a waste of time" to "always of much value," could produce total responses that mix quantity with quality measures.[21] Similarly,

---

[19]James N. McClelland, "The Effect of Student Evaluations of College Instruction Upon Subsequent Evaluations," California Journal of Educational Research, 31 (March, 1970), 89.

[20]Richard L. Turner, et al., "How do Student Characteristics Affect Their Evaluations of Instructors," Indiana University School of Education Bulletin, 45 (July, 1969), 67.

[21]C. T. Stewart and L. F. Malpass, "Estimates of Achievement and Ratings of Instructors," Journal of Educational Research, 59 (April, 1966), 347.

should "poor" matches with "outstanding" for one item, also be combined with "excellent" on another?[22] It seemed clear that by comparison, if the Critical Incident and factor analytic techniques provided highly correlated instrument items, there would be a need for more uniform scaling techniques.

One scaling technique already employed by two recent investigators was the Osgood Semantic Differential Technique.[23] Its bipolar pairs of adjectives had a high degree of reliability and validity with college students and was therefore useful for the items in the current study.[24] By employing a seven-point scale, moreover, the semantic differential permitted greater amplitude at the extremes, and provided the mid-point position as well. The only substantive criticism raised concerning the scale had been the usual one regarding the neutrality of the mid-point. Yet it appeared, from recent evidence, that

---

[22]Ibid. See also Richard R. Renner, "A Successful Rating Scale," Improving College and University Teaching, 15 (Winter, 1967), 13.

[23]Charles E. Osgood, George J. Suci, and Percy H. Tannenbaum, The Measurement of Meaning (Urbana, Illinois: University of Illinois Press, 1957) and James G. Snider and Charles E. Osgood (eds.), Semantic Differential Technique: A Sourcebook (Chicago, Aldine Publishing Company, 1969).

[24]James F. Brinton, "Deriving an Attitude Scale from Semantic Differential Data," Public Opinion Quarterly, 25 (Summer, 1961), 289-295 and Lawrence A. Pervin, "Satisfaction and Perceived Self-Environment Similarity: A Semantic Differential Study of Student-College Interaction," Journal of Personality, 35 (December, 1967), 623-634.

that this did not impair the usefulness of the semantic differential.[25]

Despite their dissimilar designs, the studies by Lewis and Rees presented promising applications of the semantic differential to evaluations studies. Lewis used the bipolar adjectives to determine whether college professors were perceived in uniformly stereotypic terms, and found that different disciplines evoked significantly different images in students.[26] Rees's more elaborate investigation utilized the semantic differential on an evaluations instrument based on the Michigan factor analysis studies.[27] The instrument was administered along with a number of standardized personality measurements to determine which characteristics might be associated with certain ratings. He used an ungrouped sample of 65 students from a sectarian university, and found that lower socioeconomic status students from stable homes presented the least favorable image of college instructors, viewing them as "generally unskilled in teaching ability."[28] Despite the limited conclusion, the Rees study not only extended the use of the

---

[25]E. R. Oetting, "The Effect of Forcing Response On the Semantic Differential," Educational and Psychological Measurement, 67 (Autumn, 1967), 699-702.

[26]Lionel S. Lewis, "Students' Images of Professors," Educational Forum, 32 (January, 1968), 189.

[27]Richard D. Rees, "Dimensions of Students Points of View in Rating College Teachers," Journal of Educational Psychology, 60 (December, 1969), 476-482.

[28]Ibid., p. 481.

sematic differential, but also represented the most comprehensive attempt, since Riley, to integrate student input characteristics with their images of instructors.[29]

After adopting the semantic differential technique for scaling purposes, the second problem was to select the appropriate bipolar adjectives for each item. Since the bipolar adjectives of the semantic differential had acquired widespread use and validation, this study selected only one bipolar pair per instrument item. With necessary modifications being made for the two instruments, both the expectations and the evaluations items used the same paired adjectives. Both instruments are appended. The instrument items presented below for expectations were classified by Critical Incident (hereafter CI) categories. It should be noted that in all but the first category, which covered the widest range of behaviors, there were three items per group. Items were identified by their instrument part and number; the semantic differential paired adjectives are noted in parentheses:

CI Category I.--Presentation of Material-Content, Structure and Scope

Part I, Item 3. What do you expect the instructor's knowledge of the subject to be? (Superior, Inferior)

Part II, Item 1. What kind of understanding of the course objectives do you expect to have? (Hazy, Clear)

---

[29]John W. Riley, Jr., Bryce F. Ryan, and Marcia Lifshitz, The Student Looks At His Teacher (New Brunswick, New Jersey: Rutgers University Press, 1950).

Part II, Item 7. Disregarding the instructor and the way the course is taught, how would you rate the subject matter of the course? (Valuable, Worthless)

Part II, Item 10. Considering everything, how would you expect to rate this course? (Good, Bad)

CI Category II.--Presentation of Material: Student Participation

Part I, Item 7. What kind of emphasis do you expect the instructor to place on the stimulation of thinking and ideas? (Heavy, Light)

Part I, Item 8. What percentage of students do you expect to finish the required reading prior to class? (Small, Large)

Part II, Item 6. How do you expect this course to relate to your own life? (Low, High)

CI Category III.--Presentation of Material: Instructor's Style

Part I, Item 2. Do you expect the instructor to make clear how each topic fits into the total course? (Actively, Passively)

Part I, Item 6. Do you expect to find the instructor using enough examples and explanations to clarify the material? (Sufficient, Insufficient)

Part I, Item 10. Considering everything, how would you recommend this instructor to a friend? (Bad, Good)

CI Category IV.--Teacher-Student Rapport and Class Interaction

Part I, Item 1. If students request help, in what manner do you expect the instructor to respond? (Pleasant, Unpleasant)

Part I, Item 4. How do you expect the instructor to react to student questions, disagreements or expressions of their own ideas? (Approving, Disapproving)

Part I, Item 5. To what extent do you expect the instructor to maintain student interest in the course? (Low, High)

CI Category V.--Evaluation of Students

Part I, Item 9. What kind of emphasis do you expect the instructor to place on accumulating factual information? (Strong, Weak)

Part II, Item 5. How do you expect to find the content of the tests and exams? (Unfair, Fair)

Part II, Item 8. In thinking about your grade, how do you expect it to reflect your true worth in the course? (Fair, Unfair)

CI Category VI.--Requirements of Students

Part II, Item 2. In relation to your other courses carrying equal credit, how do you expect the amount of study and preparation time to compare? (Just, Unjust)

Part II, Item 3. How do you expect to find the content of the assigned reading? (Worthless, Valuable)

Part III, Item 4. What kind of challenge do you expect this course to be for you? (Superior, Inferior)

The twentieth item sought to ascertain student appraisal of both the expectations and evaluations instruments. This measure, usually absent from evaluations instruments, read: How do you rate the ability of this survey to determine your own expectations /~evaluations~/ for this course? (Strong, Weak).

Hypotheses and Questions Considered

The major hypothesis confronted the notion that student's remained relatively inert in weighing their actual classroom experience, that:

One, within any class there would be no statistically significant

difference in student expectations in the six Critical Incident categories and the subsequent student evaluations.

Year in college was treated as a separate set of hypotheses in order to focus directly on the questions raised by Stern regarding the so-called "Freshman Myth." That:

Two, freshmen would not exhibit statistically significant higher contextual expectations than other classmen in any one of the six Critical Incident categories relating to the instructor, and,

Three, freshmen would not exhibit statistically significant higher evaluations compared with other classmen in any one of the six Critical Incident categories, and,

Four, there would be no statistically significant difference between freshmen expectations and evaluations compared with sophomores in any one of the six Critical Incident categories.

Raising this question regarding the influence of year in college on expectations and evaluations led to a consideration of other student variables. Accordingly, the expectations schedule included items (see Appendix) reporting certain demographic information regarding the student population. The relative association of these variables was tested in the following hypotheses, that:

Five, there would be no statistically significant difference in each of the Critical Incident expectations categories for these variables: sex, type of secondary school attended, size of secondary

school, size of residence city, social science or non-social science major, previous instructor ratings experience, and whether the course was required or elective, and,

Six, there would be no statistically significant difference in each of the Critical Incident evaluations categories for these variables: sex, type and size of secondary school attended, size of residence city, social science or non-social science major, previous instructor ratings experience, and whether the course was required or elective.

This study also sought to examine the relationship of certain student values with particular expectations or evaluations categories. So in addition to administering the expectations and evaluations instruments, the Allport-Vernon-Lindzey (AVL) Study of Values was taken by the student sample.[30] The recent Rokeach Value Study was not selected because two participating instructors mentioned introducing it as an example in a class discussion during the Spring 1970.[31] The AVL lent itself well to the current study for it amply demonstrated that the student value domain, especially for freshmen remained relatively constant.[32] Thus, the use of the AVL permitted expectations

---

[30]Gordon W. Allport, Philip E. Vernon and Gardner Lindzey, Study of Values (Boston: Houghton Mifflin Company, 1960).

[31]Milton Rokeach, "The Role of Values in Public Opinion Research," Public Opinion Quarterly, 32 (December, 1968), 548-559.

[32]Walter T. Plant and Charles W. Telford, "Changes in Personality for Groups Completing Different Amounts of College Over Two Years," Genetic Psychology Monographs, 74 (August, 1966), 3-36.

and evaluations to be compared with a presumably stable variable. The AVL itself measured the relative strength of six values or interest modes for an individual, and the means for groups indicate the relative strength of the values of the groups in question. To this point, would the values measured by the AVL (Theoretical, Economic, Aesthetic, Social, Political and Religious) produce similar or different expectations? For example, would students having high expectations for content and structure, as compared with teacher rapport, yield similar AVL values? In short, would the AVL values behave independently of the various Critical Incident categories?

CHAPTER IV

PRESENTATION OF FINDINGS

The findings have been organized and discussed according to the hypotheses and questions posed by this study. Since the data will be presented in terms of the Critical Incident (CI) categories for the study's expectations and evaluations instruments, they should be restated here: Category I--Presentation of Material: Content, Structure and Scope will be referred to as Content; Category II--Presentation of Material: Student Participation will be referred to as Participation; Category III--Presentation of Material: Instructor's Style will be referred to as Style; Category IV--Teacher-Student Rapport and Class Interaction will be referred to as Rapport; Category V--Evaluation of Students will be referred to as Evaluation; and, finally, Category VI--Requirements of Students will be referred to as Requirements. Likewise, it should be mentioned that Alpha is a two-year women's college, Beta is a comprehensive community college, and Gamma is a four-year women's college recently turned co-educational. The four Alpha College classes will be cited as A, B, C and D; the two Beta College classes as E and F; and, the two Gamma College classes as G and H.

All analyses of reported data were carried out employing an IBM Systems 1130 Computer. Conventional statistical procedures were used for testing the hypotheses and discussing the questions examined by this study. Statistically significant differences were accepted at the .05 level of judgment. For each of the hypotheses, two-tailed tests were used because the null hypotheses did not include a prediction of direction of change. The results of the study will now be presented in order of the stated hypotheses.

Hypotheses Considered

The first--and major--hypothesis considered the changes in each sample class. To test the null hypothesis of no differences between expectations and evaluations for each CI category, group means were compared using the t test between correlated means.[1] In each case degrees of freedom were determined by the number of test pairs minus one. The relevant data can be found in Tables 1, 2 and 3. In Beta and Gamma Colleges significant differences were found in virtually all of the categories in each of the classes and college totals. Most of the differences were at the .01 level. Alpha College classes reported some significant differences in three of the four classes. The College

---

[1]John H. Mueller, Karl F. Schuessler, Herbert L. Costner, Statistical Reasoning in Sociology (2d ed.; Boston: Houghton Mifflin and Company, 1970), p. 417.

TABLE 1.--Means and Standard Deviations for Expectations and Evaluations Compared by CI Category for the four classes in Alpha College and College total

Class and College		CI Category												
		Content		Participation		Style		Rapport		Evaluation		Requirements		
		$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	(N)
A	Ex.	26.1	.8	17.5	2.4	19.5	.3	19.6	.9	16.2	2.3	18.8	1.0	(23)
	Ev.	25.0	1.6	17.9[a]	1.7	18.8	1.8	19.2	1.4	16.9	2.6	17.5[a]	1.5	
B	Ex.	24.8	2.2	16.5	2.5	18.0	2.5	19.4	1.1	16.2	2.7	17.9	1.6	(21)
	Ev.	25.6	1.5	16.3	1.7	18.8	1.2	18.7	1.6	17.7	1.9	18.1	1.8	
C	Ex.	25.2	1.5	17.0	1.6	18.7	1.3	19.7	.7	16.7	2.9	18.2	1.5	(29)
	Ev.	24.3[a]	1.4	15.5[a]	2.9	18.6	1.6	18.9	1.6	16.5	2.3	17.1[a]	2.1	
D	Ex.	24.6	2.2	17.3	1.6	19.0[b]	1.5	19.9	.9	17.0	1.7	17.9[a]	2.2	(23)
	Ev.	23.0	3.6	14.2[b]	3.3	16.5[b]	3.7	17.9[b]	2.1	16.5	1.9	16.6[a]	1.8	
Alpha Total														
	Ex.	25.2	1.8	17.1	2.1	18.8	1.6	19.7	.9	16.6	2.5	18.2	1.9	(96)
	Ev.	24.4[a]	2.4	15.9[b]	2.9	18.2	2.4	18.7[a]	1.8	16.8	2.3	17.3[a]	1.9	

[a] t ratios significant at .05 level.

[b] t ratios significant at .01 level.

TABLE 2.--Means and Standard Deviations for Expectations and Evaluations Compared by CI Category for the two classes in Beta College and College total

Class and College		CI Category												
		Content		Participation		Style		Rapport		Evaluation		Requirements		
		$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	(N)
E	Ex.	23.4	2.6	15.3	2.9	17.8[b]	2.2	19.2[b]	2.0	14.4[b]	2.8	16.2[b]	2.5	(25)
	Ev.	18.6[b]	4.5	11.3[b]	3.5	11.0	4.5	11.7	4.2	10.5	3.8	12.6[b]	3.6	
F	Ex.	22.9[b]	2.4	15.5[b]	2.0	17.9[b]	.8	18.5[b]	1.2	16.2[b]	.7	16.8[b]	1.5	(24)
	Ev.	18.6	3.2	13.0[b]	2.3	14.5[b]	1.4	15.5[b]	1.3	12.8[b]	2.2	13.0[b]	2.4	
Total														
	Ex.	23.2[b]	2.5	15.4	2.5	17.9[b]	1.6	18.9[b]	1.7	15.3[b]	2.3	16.5	2.1	(49)
	Ev.	18.6[b]	3.9	12.1[b]	3.1	13.2[b]	3.6	13.5[b]	3.7	11.6	3.3	12.8[b]	3.1	

[a] t ratios significant at .05 level.

[b] t ratios significant at .01 level.

TABLE 3.--Means and Standard Deviations for Expectations and Evaluations Compared by CI Category for the two classes in Gamma College and College total

Class and College		CI Category												
		Content		Participation		Style		Rapport		Evaluation		Requirements		
		$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	(N)
G	Ex.	24.9	1.8	16.3[a]	2.2	18.1	2.3	19.5[b]	1.9	15.1[a]	2.0	17.6[b]	1.7	(40)
	Ev.	23.9[a]	2.4	15.3[a]	2.6	17.7	2.5	18.6[b]	1.7	13.8[a]	2.4	16.4[b]	2.9	
H	Ex.	24.0	1.3	14.6[b]	3.4	15.9[b]	3.2	18.2[b]	2.5	15.2	2.7	17.6[b]	2.2	(24)
	Ev.	21.0[b]	3.0	12.1[b]	2.9	13.2	3.8	13.6	3.0	13.5	2.8	14.4[b]	2.3	
Gamma Total														
	Ex.	24.6[b]	1.7	15.7[b]	2.8	17.3[b]	2.9	19.0[b]	2.2	15.1[b]	2.3	17.6[b]	1.9	(64)
	Ev.	22.8[b]	3.0	14.1[b]	3.1	16.0[b]	3.7	16.7[b]	3.7	13.7[b]	2.6	15.6[b]	2.9	

[a] t ratios significant at .05 level.

[b] t ratios significant at .01 level.

total produced significant differences in four of the six CI categories. In all, 30 out of 48 expectations-evaluations CI comparisons reached acceptable levels of statistical significance. In only 4 instances in Classes A and B did the mean differences for the evaluations outweigh the expectations; of these one instance produced a significant difference in the Participation category of Class A. Thus, the null hypothesis was rejected for the alternative that differences were to be found between student contextual expectations and evaluations.

An analysis of the ranks for the expectations and evaluations categories confirmed the alternative hypothesis. This analysis might help answer questions raised regarding the reliability of the study population. Data reported in Table 4 indicated the rank order (rho)[2] computations for the CI expectations scales in each of the colleges. Not only did the data support previous ideal professor studies noting that social science students consider instructor Rapport, course Content and Style the more important course attributes, but the rank differences between the colleges were negligible. Furthermore, the relative constancy of the students' expectations can be seen by comparing the rank order (rho) differences of their expectations and their subsequent evaluations. Table 5 reported this data. Despite the varying group mean differences actually recorded by each class, the students ranked

---

[2]John H. Mueller, Karl F. Schuessler, Herbert L. Costner, Statistical Reasoning in Sociology (2d ed.; Boston: Houghton Mifflin Company, 1970), p. 273.

TABLE 4.--Comparison of Rank Differences by CI Expectations Category for Alpha, Beta and Gamma Colleges

CI Category	College, Rank and Mean					
	Alpha[a]		Beta[b]		Gamma	
	Ex.		Ex.		Ex.	
	Rank	$\overline{X}$	Rank	$\overline{X}$	Rank	$\overline{X}$
Content	2	(6.3)	3	(5.8)	2	(6.2)
Participation	5	(5.7)	6	(5.0)	5	(5.2)
Style	3	(6.2)	2	(5.9)	3	(5.8)
Rapport	1	(6.5)	1	(6.3)	1	(6.3)
Evaluation	6	(5.5)	5	(5.1)	6	(5.0)
Requirement	4	(6.0)	4	(5.5)	4	(5.7)

[a]Rho=1.0 with Gamma College.

[b]Rho=.90 with Alpha and Gamma Colleges.

TABLE 5.--Comparison of Rank Differences by CI Expectations Category with CI Evaluations Category for Alpha, Beta and Gamma Colleges

CI Category	College, Rank and Mean					
	Alpha[a]		Beta[b]		Gamma[c]	
	Ex.	Ev.	Ex.	Ev.	Ex.	Ev.
	Rank $\overline{X}$	Rank $\overline{X}$	Rank $\overline{X}$	Rank $\overline{X}$	Rank $\overline{X}$	Rank $\overline{X}$
Content	2 (6.3)	2.5 (6.1)	3 (5.8)	1 (4.6)	2 (6.2)	1 (5.7)
Participation	5 (5.7)	6 (5.3)	6 (5.0)	5 (4.0)	5 (5.2)	5 (4.7)
Style	3 (6.2)	2.5 (6.1)	2 (5.9)	3 ((4.4)	3 (5.8)	3 (5.8)
Rapport	1 (6.5)	1 (6.2)	1 (6.3)	2 (4.5)	1 (6.3)	2 (5.5)
Evaluation	6 (5.5)	5 (5.6)	5 (5.1)	6 (3.8)	6 (5.0)	6 (4.6)
Requirement	4 (6.0)	4 (6.0)	4 (5.5)	4 (4.2)	4 (5.7)	4 (5.2)

[a]Rho=.93.

[b]Rho=.76.

[c]Rho=.94.

the six CI categories in a rather consistent fashion. This seemed remarkable even if one were faced with the fact that the students' expectations were meet with uneven outcomes. The rho ratios for the CI categories for Alpha College was .93, for Beta .76, and for Gamma .94. Certainly this would serve as a confirmation of the substantive hypothesis that expectations influence the way students record their evaluations of instructors.

This conclusion seemed to be supported by noting student reactions to the expectations and evaluations instruments. Each instrument contained an item (Part II, Item 9) eliciting student response for the ability to determine their expectations and evaluations. This data by sex can be found in Table 6, with a t test for correlated means computed for differences. The only difference actually noted indicated that the evaluations instrument proved to be a better indicator of student response. Again, despite the significant differences between actual expectations and evaluations, student response to the study instruments remained relatively constant.

To test the second hypothesis that there would be no difference between freshmen and other classmen group means for each CI expectations category were computed and reported in Table 7. Because there were so few third and fourth year students in each introductory social science course, the data were grouped as freshmen and sophomores. The t test

TABLE 6.--Means and Standard Deviations of Student Responses to Part II., Item 9 on the Evaluations Instruments for Alpha, Beta and Gamma Colleges by Sex

College	Female			Male		
	$\overline{X}$	SD	(N)	$\overline{X}$	SD	(N)
Alpha						
Ex.	5.49	.94	(96)	--	--	--
Ev.	5.88[a]	.64		--	--	--
Beta						
Ex.	5.40	.76	(25)	5.46	.72	(24)
Ev.	5.48	.77		5.50	.83	
Gamma						
Ex.	5.82	.65	(45)	5.84	.60	(19)
Ev.	5.87	.51		5.79	.54	

[a] t ratios significant at .05 level.

TABLE 7.--Mean and Standard Deviation Comparisons of Freshmen with Sophomores by CI Expectation Category within Alpha, Beta, and Gamma Colleges

College and Year	CI Category												
	Content		Participation		Style		Rapport		Evaluation		Requirements		
	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	(N)
Alpha													
Fresh.	25.0	2.0	17.4	1.9	19.2	1.2	19.7	.7	16.9	2.6	18.2	2.0	(44)
Soph.	25.4	1.6	16.7	2.2	18.5	1.9	19.6	1.0	16.2	2.3	18.3	1.7	(52)
Beta													
Fresh.	23.1	2.3	15.4	2.8	17.7	2.0	18.7	2.0	15.1	2.3	16.3	2.3	(30)
Soph.	23.2	2.7	15.5	2.0	18.2	1.0	19.2	1.1	15.6	2.1	16.7	1.7	(19)
Gamma													
Fresh.	24.3	1.5	15.5	3.0	17.2	3.0	19.0	1.7	15.5	2.0	17.4	2.3	(29)
Soph.	24.8	1.9	15.8	2.7	17.4	2.8	19.0	2.6	15.2	2.6	17.7	1.4	(35)

for uncorrelated--or independent--means[3] was used for determining statistically significant differences. Degrees of freedom were determined by summing the two test samples minus two. No differences were noted. The null hypothesis was therefore accepted. In this case the null hypothesis served as the substantive one. Clearly, the freshmen and sophomores reported similar contextual expectations. Even if one discounts the fact that students at the two-year colleges would be subject to an "upperclassmen" effect, no differences were reported by the four-year college sophomores.

The third hypothesis that there would be no differences in freshman and sophomore evaluations was similarly tested and cited in Table 8. As can be readily seen, the only differences noted were for sophomores--and those were significantly greater. Although sophomore evaluations were significantly higher in nine instances, in all but one their evaluations were still higher than freshmen. Here again the null hypothesis was accepted, for when differences appeared sophomore evaluations--not freshmen--produced them. Comment on this can be reserved following the consideration of the next hypothesis.

Hypothesis four was tested on the point that there would be no significant differences between freshmen expectations and evaluations

---

[3]John H. Mueller, Karl F. Schuessler, Herbert L. Costner, Statistical Reasoning in Sociology (2d ed.; Boston: Houghton Mifflin Company, 1970), p. 407.

TABLE 8.--Mean and Standard Deviation Comparisons of Freshmen with Sophomores by CI Evaluations Category within Alpha, Beta, and Gamma Colleges

College and Year	CI Category												
	Content		Participation		Style		Rapport		Evaluation		Requirements		
	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	(N)
Alpha													
Fresh.	23.9	2.7	15.6	3.0	17.8	2.8	18.4	2.5	16.8	2.5	17.3	2.0	(44)
Soph.	25.0[a]	1.9	16.1	2.7	18.5	2.0	19.0	1.5	16.9	2.0	17.3	1.9	(52)
Beta													
Fresh.	18.3[b]	3.9	11.6[b]	2.6	12.6[c]	3.4	12.8	3.3	11.4	3.3	12.1	3.4	(30)
Soph.	19.2[b]	3.8	13.0[b]	3.7	14.1[c]	3.6	14.6[c]	3.9	11.9[a]	3.3	13.9[c]	2.2	(19)
Gamma													
Fresh.	22.6	2.6	13.8	2.8	14.7	4.3[b]	15.8[a]	3.6	13.7	2.5	15.4	2.7	(29)
Soph.	23.0	3.3	14.3	3.2	17.1	2.7[b]	17.4[a]	3.0	13.7	2.6	15.7	3.0	(35)

[a] $\underline{t}$ ratios significant at .05 level.

[b] $\underline{t}$ ratios significant at .01 level.

[c] $\underline{t}$ ratios significant at .001 level.

as compared to sophomores. The group means for this comparison by college can be found in Table 9. Here the t test for correlated means was employed. Since there were significant differences to the .001 level for both freshman and sophomore evaluations the null hypothesis became untenable. This would confirm the substantive hypothesis of this study--students record their classroom experiences in similar ways. Taking Beta College, for example, it did not make much difference whether freshmen had lower expectations than sophomores, the actual experiences produced significantly lower evaluations for both groups. In summing up hypotheses three, four and five, both freshmen and sophomores in this study recorded expectations and evaluations that seemed to relate to each other's actual experiences rather than their year in college.

Because the fifth and sixth hypotheses dealt with the differences produced by certain variables in relation to expectations and evaluations, each variable will be discussed independently. Each hypothesis stated that no differences would be exhibited in any CI Category for either expectations or evaluations with the following demographic variables: type of secondary school attended, size of residence city, social science or non-social science major, previous evaluations experience, elective or required status of the course, size of secondary school graduating class, and sex. Each of the seven variables was tested for expectations and evaluations by college employing an uncorrelated t test. For separate group variance between groups of unequal size, degrees of freedom were determined at the mid-point between the paired samples minus one for each

TABLE 9.--Means and Standard Deviations Compared by CI Expectations and Evaluations Category for Year in Alpha, Beta and Gamma Colleges

College and Year	CI Category												
	Content		Participation		Style		Rapport		Evaluation		Requirements		
	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	(N)
Alpha													
Fr. Ex.	25.0	2.0	17.4	1.9	19.2	1.2	19.7[b]	2.6	16.8	2.6	18.2	2.0	(44)
Ev.	23.8[a]	2.7	15.8[b]	3.0	17.8[a]	2.8	18.4	2.5	16.8	2.5	17.3[a]	2.0	
So. Ex.	25.4	1.6	16.7	2.2	18.5	1.9	19.6	1.0	16.2	2.3	18.3[a]	1.7	(52)
Ev.	25.0	1.9	16.1	2.7	18.5	2.0	19.0	1.5	16.9	2.0	17.3	1.9	
Beta													
Fr. Ex.	23.1	2.3	15.4	2.8	17.7	2.0	18.7[c]	2.0	15.1[c]	2.3	16.3	2.3	(30)
Ev.	18.3[c]	3.9	11.6[c]	2.6	12.6[c]	3.4	12.8[c]	3.3	11.4	3.3	12.1[c]	2.2	
So. Ex.	23.3	2.7	15.5	2.0	18.2	1.0	19.2[c]	1.1	15.6	2.1	16.7[c]	1.7	(19)
Ev.	19.2[c]	3.8	13.0[a]	3.7	14.1[c]	3.6	14.6[c]	3.9	11.9[c]	3.3	13.9[c]	3.4	
Gamma													
Fr. Ex.	24.3[b]	1.5	15.5	3.0	17.2[b]	3.0	19.0	1.7	15.5	2.0	17.4	2.3	(29)
Ev.	22.6	2.6	13.8[a]	2.8	14.7	4.3	15.8[c]	3.6	13.7[a]	2.6	15.4[c]	2.7	
So. Ex.	24.8[b]	1.9	15.8[b]	2.7	17.4	2.8	19.0[b]	2.6	15.2	2.5	17.7	1.4	(35)
Ev.	23.0	3.3	14.3[b]	3.2	17.1	2.7	17.4[b]	3.0	13.7[b]	2.6	15.7[b]	3.0	

[a] $\underline{t}$ ratios significant at .05 level.

[b] $\underline{t}$ ratios significant at .01 level.

[c] $\underline{t}$ ratios significant at .001 level.

pair. The data appear in Tables 10 to 16.

The comparisons for public with private secondary school graduates for each expectations and evaluations category is reported in Table 10. In only one instance--Gamma College--did attendance at either public or private school show significantly different expectations. The finding for this variable would contribute to accepting the null hypotheses--that is, the substantive ones--of no differences between types of school attended.

Table 11 presents the data comparing students' residence city size for expectations and evaluations. The two categories of small and large city size, noted in the table, resulted in a bimodal distribution of responses to the four ranges presented in the expectations instrument (see Appendix A). Most students in the three colleges fell into either small suburban areas with populations of 50,000 or fewer people, and medium to somewhat larger size cities of 50,000 to 250,000 people. The data was organized according to this distribution in order to develop a sample size that would enhance differences. Remarkably, no differences emerged--the only demographic variable to produce such results. This would confirm the null hypotheses, even though the population lacked students from large cities.

The data for social science majors compared with non-social science majors appear in Table 12. As can be readily seen, Beta College recorded the only statistically significant difference for an expectations category. This evidence would contribute to accepting the null hypotheses.

TABLE 10.--Means and Standard Deviations by Expectations and Evaluations CI Category Comparing Public with Private Secondary School Graduates according to Alpha, Beta and Gamma Colleges

College		CI Category and School Type												
		Content		Partici-pation		Style		Rapport		Evalua-tion		Require-ments		
		Pb.	Pr.	Pb.	Pr.	Pb.	Pr.	Pb.	Pr.	Pb.	Pr.	Pb.	Pr.	(N)
Alpha														
Ex.	$\overline{X}$	25.0	25.3	17.1	17.0	19.0	18.6	19.7	19.7	16.3	16.8	18.4	18.0	Pb.
	SD	2.2	2.0	2.4	2.3	1.8	2.0	1.3	1.4	2.5	2.9	2.2	2.1	(46)
Ev.	$\overline{X}$	24.3	24.5	15.8	16.0	17.8	18.6	18.3	19.0	16.8	16.9	17.7	17.0	Pr.
	SD	2.8	2.3	3.0	3.1	3.2	1.9	2.2	1.8	2.5	2.5	1.7	2.5	(50)
Beta														
Ex.	$\overline{X}$	23.1	23.8	15.2	17.2	18.0	17.0	18.8	20.0	15.4	14.4	16.5	16.0	Pb.
	SD	2.7	3.1	2.0	1.6	2.0	1.6	2.0	1.0	2.5	2.4	2.3	3.3	(44)
Ev.	$\overline{X}$	18.4	20.6	12.4	10.2	13.4	10.8	13.6	13.4	11.8	10.0	12.9	12.2	Pr.
	SD	4.1	4.0	3.3	3.6	3.5	5.3	3.8	4.6	3.5	3.8	3.1	5.3	(5)
Gamma														
Ex.	$\overline{X}$	24.2	25.2	15.7	15.7	17.0	17.8	19.0	19.0	15.0	15.4	17.5	17.7	Pb.
	SD	2.2	1.4	3.0	3.2	3.4	2.5	2.6	2.2	2.8	2.1	2.4	1.8	(40)
Ev.	$\overline{X}$	22.4	23.4	13.3	15.4[a]	15.3	17.1	16.2	17.5	13.2	14.5	15.1	16.5	Pr.
	SD	3.4	2.7	3.4	2.5	4.3	3.0	3.9	2.7	3.1	2.0	3.1	2.7	(24)

[a] $\underline{t}$ ratios significant at .05 level.

[b] $\underline{t}$ ratios significant at .01 level.

TABLE 11.--Means and Standard Deviations by Expectations and Evaluations CI Category Comparing Students by Size of Residence City according to Alpha, Beta and Gamma Colleges

College		CI Category and City Size												
		Content		Participation		Style		Rapport		Evaluation		Requirements		
		Sm.	Lg.	Sm.	Lg.	Sm.	Lg.	Sm.	Lg.	Sm.	Lg.	Sm.	Lg.	(N)
Alpha														
Ex.	$\bar{X}$	25.0	25.6	16.9	17.3	18.8	18.8	19.7	19.6	16.3	17.0	18.1	18.4	Sm.
	SD	2.2	1.8	2.4	2.3	2.1	1.7	1.4	1.5	2.8	2.6	2.2	2.1	(62)
Ev.	$\bar{X}$	24.5	24.2	16.1	15.8	18.2	18.3	18.6	18.9	16.9	16.8	17.5	16.9	Lg.
	SD	2.4	3.0	3.0	3.2	2.7	2.7	2.2	1.6	2.6	2.3	1.9	2.6	(34)
Beta														
Ex.	$\bar{X}$	23.0	23.6	15.5	15.3	17.9	17.7	18.8	19.1	15.3	15.2	16.7	15.9	Sm.
	SD	2.9	2.1	2.5	3.2	1.4	3.0	1.9	2.3	2.2	3.1	2.3	2.5	(35)
Ev.	$\bar{X}$	18.4	19.2	11.8	12.9	13.2	13.0	13.3	14.0	11.7	11.6	13.0	12.2	Lg.
	SD	4.0	4.5	2.7	4.5	3.6	4.1	3.7	4.4	3.5	3.6	3.3	3.3	(14)
Gamma														
Ex.	$\bar{X}$	24.5	24.6	15.3	15.9	16.5	18.0	18.8	19.2	15.0	15.2	17.4	17.8	Sm.
	SD	2.1	1.9	3.6	2.4	3.6	2.4	2.6	2.4	2.9	2.2	2.1	2.3	(29)
Ev.	$\bar{X}$	23.1	22.6	14.3	13.9	15.9	16.0	16.4	16.9	13.4	13.9	16.2	15.1	Lg.
	SD	3.3	3.2	3.4	3.2	4.3	3.6	4.0	3.1	2.6	2.9	2.5	3.4	(35)

TABLE 12.--Means and Standard Deviations by Expectations and Evaluations CI Category Comparing Students by Social Science Majors with Non-Social Science Majors According to Alpha, Beta and Gamma Colleges

College		CI Category and Major												
		Content		Participation		Style		Rapport		Evaluation		Requirements		
		Maj.	Non.	Maj.	Non.	Maj.	Non.	Maj.	Non.	Maj.	Non.	Maj.	Non.	(N)
Alpha														
Ex.	$\bar{X}$	25.6	25.0	17.8	16.6	19.3	18.9	19.3	19.6	17.2	16.2	18.4	17.8	Maj.
	SD	1.8	2.2	2.3	2.4	1.7	1.5	1.3	1.5	2.1	3.0	2.4	2.2	(23)
Ev.	$\bar{X}$	24.5	24.6	16.5	16.0	18.1	18.1	18.8	18.8	16.5	17.0	17.6	17.1	Non.
	SD	2.9	2.2	2.8	3.1	2.5	2.5	2.0	1.7	2.7	2.4	1.8	2.3	(42)
Beta														
Ex.	$\bar{X}$	24.2	22.9	18.0	14.9[a]	17.6	17.9	19.8	18.9	15.8	15.4	18.0	16.5	Maj.
	SD	3.4	2.7	2.9	2.6	.9	2.2	.8	2.1	2.4	2.4	1.0	2.3	(5)
Ev.	$\bar{X}$	18.4	18.6	11.2	12.1	10.8	13.2	13.2	13.5	11.8	11.4	12.2	13.2	Non.
	SD	4.2	4.4	4.4	3.3	4.9	3.9	3.6	4.2	4.3	3.5	5.1	3.2	(35)
Gamma														
Ex.	$\bar{X}$	24.4	25.4	15.2	15.9	17.0	16.7	18.9	19.0	15.5	14.6	17.7	17.0	Maj.
	SD	1.7	2.2	3.3	3.8	3.2	2.8	2.2	3.1	2.5	3.5	2.0	2.0	(30)
Ev.	$\bar{X}$	22.2	22.9	13.5	13.1	14.7	17.2	15.5	17.2	13.8	14.2	15.4	15.7	Non.
	SD	3.3	3.4	3.4	2.8	4.4	2.3	3.9	3.3	3.0	2.7	2.6	3.2	(9)

[a] t ratio significant at .05 level.

Table 13 discloses the data relating the influence of previous instructor rating experience on expectations and evaluations. Gamma College showed the only significant differences, one for expectations and the other for an evaluations category. This, too, would not lend any substantial evidence for rejecting the null hypotheses.

The comparison of students taking the course for elective or required credit appears in Table 14. Gamma College reported the only significant differences, with two each for the expectations and for the evaluations categories. Again, this evidence would be insufficient to reject the null hypotheses.

Table 15 shows the data comparing students by their secondary school graduating class size. Because the responses to the four categories noted in the expectations instrument (see Appendix A) again produced a bimodal distribution of small--less than 200 students in a graduating class--compared with large--more than 200 in a class--the data were organized around these categories. Alpha College presented no statistically significant differences, Beta College revealed only one for an expectations category.

The responses from Gamma College, however, indicated that in four of the six evaluations categories, and one expectations category, significant differences were exhibited. The only other variable to produce as many differences as secondary class size was sex, as is shown by an examination of Table 16 which reports the data for sex differences in Beta and Gamma Colleges--the co-educational ones.

TABLE 13.--Means and Standard Deviations by Expectations and Evaluations CI Category Comparing Students by Previous Instructor Rating Experience according to Alpha, Beta and Gamma Colleges

College		CI Category and Rating Experience												
		Content		Participation		Style		Rapport		Evaluation		Requirements		
		Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	(N)
Alpha														
Ex.	$\bar{X}$	25.2	25.1	17.2	16.6	18.8	18.9	19.6	19.8	16.7	16.0	18.3	18.1	Yes
	SD	1.9	2.6	2.2	2.7	1.9	2.1	1.4	1.3	2.8	2.1	2.0	2.4	(75)
Ev.	$\bar{X}$	24.6	24.0	15.9	16.0	17.9	19.1	18.6	18.9	16.9	16.8	17.3	17.3	No
	SD	2.7	2.4	3.0	3.2	2.9	1.4	2.1	1.9	2.5	2.5	2.2	2.2	(21)
Beta														
Ex.	$\bar{X}$	23.9	22.0	15.9	14.7	18.1	17.4	19.1	18.6	15.5	14.9	16.8	15.9	Yes
	SD	2.2	3.1	2.6	2.7	1.4	2.6	1.8	2.3	2.4	2.6	2.2	2.5	(30)
Ev.	$\bar{X}$	18.8	18.3	11.7	12.9	12.6	14.1	12.9	14.6	11.9	11.3	12.5	13.2	No
	SD	4.3	3.8	3.8	2.4	4.3	2.3	4.3	2.7	3.8	3.1	3.3	3.3	(19)
Gamma														
Ex.	$\bar{X}$	24.7	24.0	15.8	15.1	17.7	15.4[a]	19.2	18.0	15.2	14.9	17.6	17.9	Yes
	SD	2.1	1.8	3.0	3.5	2.9	3.6	2.1	4.0	2.5	2.9	2.2	2.2	(53)
Ev.	$\bar{X}$	23.3	20.3[b]	14.3	13.3	16.4	14.2	17.0	15.5	14.0	12.1	15.8	14.6	No
	SD	3.1	2.1	3.4	2.8	3.9	3.4	3.6	2.5	2.7	2.7	3.1	2.6	(10)

[a] $\underline{t}$ ratio significant at .05 level.

[b] $\underline{t}$ ratio significant at .01 level.

TABLE 14.--Means and Standard Deviations by Expectations and Evaluations CI Category Comparin Students by Elective with Required Course Status according to Alpha, Beta and Gamma Colleges

College		CI Category and Course Status												
		Content		Participation		Style		Rapport		Evaluation		Requirements		
		Req.	El.	Req.	El.	Req.	El.	Req.	El.	Req.	El.	Req.	El.	(N)
Alpha														
Ex.	$\bar{X}$	25.5	25.1	17.3	17.0	19.2	18.7	19.8	19.6	17.4	16.4	18.4	18.2	Req.
	SD	1.5	2.2	2.9	2.2	1.4	2.0	1.1	1.4	2.4	2.8	2.3	2.1	(17)
Ev.	$\bar{X}$	24.5	24.4	16.8	15.7	18.7	18.1	18.8	18.7	16.9	16.8	17.6	17.2	El.
	SD	2.9	2.6	2.8	3.1	1.5	2.8	2.1	2.0	2.6	2.5	2.1	2.2	(19)
Beta														
Ex.	$\bar{X}$	23.6	22.7	15.6	15.2	18.0	17.7	19.1	18.7	15.5	15.0	17.0	16.0	Req.
	SD	2.8	2.6	2.8	2.7	1.3	2.4	1.8	2.1	2.3	2.7	1.8	2.7	(25)
Ev.	$\bar{X}$	18.4	18.9	12.2	12.1	12.8	13.5	14.0	13.0	12.0	11.3	12.6	13.0	El.
	SD	3.8	4.4	3.6	3.1	3.7	3.8	3.8	3.9	3.7	3.4	3.2	3.4	(24)
Gamma														
Ex.	$\bar{X}$	24.0	25.0[a]	15.0	16.2	16.1	18.1[a]	18.5	19.3	15.6	14.8	17.3	17.8	Req.
	SD	1.8	2.0	3.4	2.6	3.3	2.6	2.5	2.4	2.8	2.3	2.3	2.0	(27)
Ev.	$\bar{X}$	21.9	23.5	13.3	14.6	14.2	17.3[b]	15.1	17.8[b]	13.5	13.8	15.0	16.0	El.
	SD	3.4	2.9	3.5	3.0	4.5	2.8	4.2	2.4	3.0	2.7	2.8	3.2	(32)

[a] $\underline{t}$ ratios significant at .05 level.

[b] $\underline{t}$ ratios significant at .01 level.

TABLE 15.--Means and Standard Deviations by Expectations and Evaluations CI Category Comparing Students by Secondary School Class Size according to Alpha, Beta and Gamma Colleges

College		CI Category and School Size												
		Content		Participation		Style		Rapport		Evaluation		Requirements		
		Sm.	Lg.	Sm.	Lg.	Sm.	Lg.	Sm.	Lg.	Sm.	Lg.	Sm.	Lg.	(N)
Alpha														
Ex.	$\overline{X}$	25.1	25.4	16.8	17.8	18.6	19.4	19.7	19.5	16.6	16.6	18.0	18.7	Sm.
	SD	2.1	2.2	2.5	1.6	2.1	1.4	1.4	1.3	3.0	1.9	2.1	2.3	(71)
Ev.	$\overline{X}$	24.4	24.4	15.9	16.1	18.3	17.9	18.9	18.0	17.0	16.5	17.1	17.9	Lg.
	SD	2.3	3.3	3.1	2.8	2.2	3.7	1.8	2.5	2.5	2.5	2.4	1.5	(25)
Beta														
Ex.	$\overline{X}$	21.8	23.7[a]	14.2	15.8	17.1	18.1	18.7	19.0	14.9	15.4	15.8	16.8	Sm.
	SD	3.5	2.2	3.2	2.4	3.0	1.3	2.7	1.7	2.3	2.6	2.8	2.1	(13)
Ev.	$\overline{X}$	18.3	18.8	12.0	12.2	13.2	13.2	14.3	13.3	11.2	11.8	13.7	12.5	Lg.
	SD	3.6	4.3	2.5	3.6	3.8	3.8	4.1	3.8	2.8	3.8	3.1	3.3	(36)
Gamma														
Ex.	$\overline{X}$	25.1	23.9[a]	15.9	15.3	17.7	16.8	19.2	18.7	15.6	14.6	17.7	17.5	Sm.
	SD	1.7	2.0	3.3	2.7	3.1	3.1	2.2	2.8	2.1	3.0	1.8	2.6	(35)
Ev.	$\overline{X}$	23.6	21.8[a]	15.1	12.9[a]	17.0	14.8[a]	17.4	15.8	14.0	13.3	16.3	14.4[b]	Lg.
	SD	3.0	3.2	3.0	3.2	3.4	4.2	2.9	4.0	2.4	3.2	2.5	3.3	(29)

[a] $\underline{t}$ ratios significant at .05 level.

[b] $\underline{t}$ ratios significant at .01 level.

TABLE 16.--Means and Standard Deviations by Expectations and Evaluations CI Category Comparing Students by Sex according to Beta and Gamma Colleges[a]

College		CI Category and Sex												
		Content		Participation		Style		Rapport		Evaluation		Requirements		
		M	F	M	F	M	F	M	F	M	F	M	F	(N)
Beta														
Ex.	$\overline{X}$	23.5	22.8	15.4	15.4	18.3	17.5	18.9	18.9	15.3	15.2	17.1	15.9	M
	SD	1.9	3.3	2.5	2.9	1.9	2.1	1.6	2.3	1.9	3.0	2.4	2.2	(24)
Ev.	$\overline{X}$	18.8	18.4	11.3	12.9	13.5	12.8	13.0	14.0	11.0	12.2	12.6	13.0	F
	SD	4.2	4.0	3.1	3.2	3.5	4.0	3.6	4.1	4.1	2.8	3.4	3.3	(25)
Gamma														
Ex.	$\overline{X}$	23.6	25.0[a]	13.8	16.5	15.7	18.0[a]	18.1	19.4	13.8	15.7[a]	15.8	18.3[b]	M
	SD	1.9	1.9	3.5	2.5	3.8	2.5	2.6	2.3	2.4	2.4	2.2	1.7	(19)
Ev.	$\overline{X}$	21.7	23.3	12.9	14.6	14.1	16.8[a]	15.5	17.1	13.3	13.8	14.1	16.2[a]	F
	SD	3.1	3.1	3.1	3.2	4.6	3.3	4.2	3.1	2.8	2.8	3.3	2.7	(45)

[a]Alpha is not a co-educational institution.

[b]$\underline{t}$ ratios significant at .05 level.

[c]$\underline{t}$ ratios significant at .01 level.

While Beta College showed no statistically significant differences when sex was compared by CI category, Gamma College evidenced almost the same pattern for differences associated with sex as it did for secondary class size. In four of the six expectations categories and two evaluations categories, females (N=45) demonstrated statistically significant higher means than male students (N=19). This evidence might appear to support the notion that the four-year college students represented an incompatible population for comparisons with the two-year colleges, or that Gamma's newly arrived men added the unusual dimension for the two variables under consideration.

The significant differences in sex and class size can be explained by examining the actual numbers of females and males in the large and small secondary school size groups in Table 15. Females outnumbered males in large secondary school by about only two to one (F=18; M=11), but they outnumbered the males three to one (F=26; M=9) in small secondary schools. Since the small school student expectations and evaluations group means significantly outweighed those of large school students, it followed that the sex difference data would show females displaying higher expectations than males. Although this evidence did not explain why the females from smaller secondary schools recorded higher expectations and evaluations, it was certainly made clear that the sex differences in this newly co-educational institution were not singularly responsible for differences in the variables noted by this study. It also confirmed what the other component variables for hypotheses five and six

domonstrated that the differences among this student population were generally slight in relation to the CI expectations and evaluations categories. Taking the foregoing into account, it did not seem to be a large risk to accept the null hypotheses for secondary school size and the sex variables. It also seemed that with the small number of significant differences reported for the other component variables, there would not be too great a risk in accepting the null hypotheses here, too. Thus, it was accepted that variables associated with the student domain would not exhibit statistically significant differences in the CI expectations and evaluations categories.

Questions Considered

The AVL Study of Values was introduced to determine how this student population compared with previous research findings on the consistency of its six scales.[4] It was also employed to learn whether certain of the CI expectations or evaluations categories might be associated with particular AVL values. The AVL group mean scores for the pretest and posttest by sex and college year can be found in Table 17 and 18. A *t* test for correlated means was computed to determine statistically significant differences. Overall the AVL values remained constant. Nonetheless, there were some significant specific changes, especially for freshmen; fifteen significant changes developed for

---

[4]They are Theoretical, Economic, Aesthetic, Social, Political and Religious.

TABLE 17.--Means and Standard Deviations for AVL Scale on Pretest and Posttest by Sex and Year In Alpha and Beta Colleges

College, Year and Sex		AVL Scale												
		Theoretical		Economic		Aesthetic		Social		Political		Religious		
	Test	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	(N)
Alpha														
Fem.	Pre-	37.23	6.3	37.57	6.9	44.48	5.4	44.66	7.6	40.50	5.9	35.20	7.2	(44)
Fresh.	Post-	38.09[c]	5.7	37.48	6.0	44.70	5.0	44.34	6.3	41.30[b]	5.0	33.86[c]	6.7	
Fem.	Pre-	37.79	6.1	35.12	6.3	46.33	5.4	43.19	6.2	41.15	7.4	36.21	7.2	(52)
Soph.	Post-	38.58[b]	5.4	35.38	5.4	46.02	4.6	43.60	5.8[b]	42.02[b]	6.6	34.60[b]	6.9	
Beta														
Male	Pre-	41.94	4.3	40.69	7.0	36.38	7.1	41.94	9.4	45.31	5.1	34.38	4.6	(16)
Fresh.	Post-	42.50[a]	3.9	40.31	6.0	36.56	6.4	42.00	7.8	45.13	4.1	33.50[c]	4.4	
Male	Pre-	42.38	3.6	40.38	4.1	40.38	4.9	42.25	5.8	42.50	3.6	32.13	3.2	(8)
Soph.	Post-	42.63	2.6	40.00	3.2	41.25	3.6	41.75	4.6	42.75	3.2	31.63	2.9	
Fem.	Pre-	34.86	4.6	36.57	6.7	43.71	7.7	45.43	8.3	37.93	6.0	41.50	8.7	(14)
Fresh.	Post-	35.93[c]	4.0	36.21	5.9	43.79	7.0	44.93	6.8	39.36[c]	5.4	40.50	7.5	
Fem.	Pre-	35.82	5.3	38.91	4.6	45.09	6.0	46.73	6.1	39.09	5.9	34.36	8.0	(11)
Soph.	Post-	37.09[c]	5.1	38.27[b]	4.1	44.90	5.3	46.27	4.7	40.00[a]	5.0	33.45[b]	7.4	

[a] $\underline{t}$ ratios significant at .05 level.

[b] $\underline{t}$ ratios significant at .01 level.

[c] $\underline{t}$ ratios significant at .001 level.

TABLE 18.--Means and Standard Deviations for AVL Scale on Pretest and Posttest by Sex and Year in Gamma College

College, Sex, and Year		AVL Scale												
		Theoretical		Economic		Aesthetic		Social		Political		Religious		
	Test	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	(N)
Gamma														
Male Fresh.	Pre-	42.07	2.7	32.07	5.4	46.13	6.2	42.13	8.4	42.47	8.5	34.47	6.8	(15)
	Post-	42.53[a]	2.7	33.33[c]	5.0	46.67	4.4	42.00	6.9	42.80	7.3	32.67[c]	6.2	
Male Soph.	Pre-	35.50	2.5	36.25	8.1	51.50	9.0	45.50	9.5	43.50	5.5	27.75	7.7	(4)
	Post-	36.50	2.5	36.00	7.5	50.75	8.2	45.25	8.2	44.25	5.0	27.25	7.9	
Fem. Fresh.	Pre-	38.57	3.3	29.79	4.6	51.36	5.7	47.07	4.0	38.79	6.2	35.14	8.2	(14)
	Post-	39.29[a]	2.7	30.79[c]	4.2	50.43[a]	5.3	47.07	3.8	40.07[c]	5.8	32.36[b]	7.2	
Fem. Soph.	Pre-	39.84	3.3	33.61	4.3	49.03	4.0	40.90	8.5	42.39	5.6	34.23	6.7	(31)
	Post-	39.84	3.6	33.65	4.3	48.81	4.0	40.94	8.2	42.32	5.8	34.45	6.8	

[a] t ratios significant at .05 level.

[b] t ratios significant at .01 level.

[c] t ratios significant at .001 level.

freshmen, and eight for sophomores. It seemed clear, however, that the general direction of these changes for both groups confirmed previous research. Change, that is, that tended to exhibit a decrease in the religious scale compared with an increase in the theoretical scale.

Were there any significant relationships between the AVL values and either the CI expectations or evaluations categories? The data by sex and college for expectations were reported in Table 19, and for evaluations in Table 20. In both cases a Pearson product-moment correlation coefficient [5] was computed for each of the items. No overall significant pattern of relationships emerged for either expectations or evaluations, however. The only pattern that developed concerned the male students at Gamma College, resulting in a pattern of inversion on the expectations scale. These students produced a significant positive correlation for the AVL political scale with an r of .69 along with content (r .01=.58) and an r of .46 for course requirements, (r .05=.46) in comparison with a negative correlation for concern with the instructor's style (r=-.56). No similar development occurred for these students on the evaluations scale. These rather inconclusive findings pointed toward additional work needed to examine college student values and contextual expectations.

---

[5]John H. Mueller, Karl F. Schuessler, Herbert L. Costner, Statistical Reasoning in Sociology (2d ed.; Boston: Houghton Mifflin Company, 1970), p. 319.

TABLE 19.--Intercorrelation Matrix of Pretest AVL Values with CI Expectations Categories by Sex for Alpha, Beta and Gamma Colleges (Decomal omitted)

College and AVL Scale	CI Category and Sex												
	Content		Partici-pation		Style		Rapport		Evalua-tion		Require-ments		
	M	F	M	F	M	F	M	F	M	F	M	F	(N)
Alpha													
Theo.	-	16	-	-10	-	-06	-	01	-	03		-05	F
Eco.	-	15	-	-21[a]	-	-05	-	15	-	-02	-	-05	(96)
Aes.	-	01	-	-21[a]	-	-13	-	21[a]	-	-07	-	14	
Soc.	-	02	-	-05	-	-02	-	05	-	-04	-	-01	
Pol.	-	-13	-	05	-	-13	-	-01	-	-21[a]	-	-03	
Rel.	-	-08	-	-13	-	-01	-	-16	-	-06	-	-07	
Beta													
Theo.	-18	-24	01	-06	04	01	33	-17	-06	12	-48[a]	23	M
Eco.	21	-11	01	03	-37	07	23	01	07	.24	10	18	(24)
Aes.	-02	14	-10	11	-17	05	24	13	-14	07	16	-25	
Soc.	33	07	-18	-13	-34	12	21	23	-11	-31	02	-03	F
Pol.	12	18	24	28	03	01	-49[a]	-22	-23	20	01	-23	(25)
Rel.	-32	26	10	15	47[a]	05	17	17	-05	05	-19	-17	
Gamma													
Theo.	14	-11	-21	-27	-11	-15	20	-07	-08	18	16	27	M
Eco.	-04	-02	-12	-17	-29	-03	-05	05	33	-08	22	12	(19)
Aes.	13	-00	-02	-02	-10	-07	11	-06	13	-07	-07	17	
Soc.	14	-01	15	-07	-32	07	06	-02	32	-12	-17	15	F
Pol.	69[b]	-16	-30	07	-56[b]	-03	-07	11	04	-02	46[a]	02	(45)
Rel.	29	-03	-15	-30[a]	-44	-10	03	18	11	-08	24	16	

[a] $\underline{r}$ significant at .05 level.
[b] $\underline{r}$ significant at .01 level.

TABLE 20.--Intercorrelation Matrix of Posttest AVL Values with CI Evaluations Categories by Sex for Alpha, Beta and Gamma Colleges (Decimal omitted)

College and AVL Scale	CI Category and Sex												
	Content		Participation		Style		Rapport		Evaluation		Requirements		
	M	F	M	F	M	F	M	F	M	F	M	F	(N)
Alpha													
Theo.		08		09		05		06		10		11	F
Eco.		05		-23[a]		14		06		-09		12	(96)
Aes.		01		17		10		10		-04		21[a]	
Soc.		11		11		12		05		-10		04	
Pol.		-11		-00		-06		03		10		04	
Rel.		-02		-01		-08		-00		-09		02	
Beta													
Theo.	-13	-13	-11	05	-30	-13	29	-21	13	17	07	17	M
Eco.	-13	24	-02	24	-09	-09	04	-07	17	15	03	-31	(24)
Aes.	03	44[a]	18	29	14	-25	23	-28	29	27	06	-33	F
Soc.	-12	40[a]	03	06	-03	-18	11	-30	05	30	-13	-21	(25)
Pol.	-22	19	36	04	24	07	-33	-55[b]	20	45[a]	-28	-16	
Rel.	-28	11	-13	22	29	16	-06	12	08	-08	-02	-31	
Gamma													
Theo.	28	-10	24	00	-43	-04	-18	-05	39	-05	-21	-10	M
Eco.	07	-09	25	-05	-31	22	04	09	28	-05	-31	-01	(19)
Aes.	-09	15	41	20	12	01	-04	11	34	-01	-48[a]	-12	F
Soc.	06	-10	17	15	-43	04	13	-16	14	-05	-10	07	(45)
Pol.	-21	-20	58[b]	-18	01	-01	-27	-05	38	-13	-46[a]	-40[b]	
Rel.	35	-03	13	04	-43	19	-30	-07	31	-03	05	-04	

[a] r significant at .05 level.
[b] r significant at .01 level.

CHAPTER V

DISCUSSION, IMPLICATIONS AND RECOMMENDATIONS

This chapter is organized in three sections. The first section will discuss and summarize the study findings. The next section presents implications of these findings. And the last section will make recommendations for related future studies. An afterword follows.

Discussion of Findings

This study examined student contextual expectations in relation to their evaluations in the natural setting of the classroom. The findings generally supported the substantive hypotheses, namely, that student evaluations were closely associated with their expectations of instructors. This was affirmed in the analyses of each class and college group means as well as rank order comparisons for the CI expectations and evaluations categories. More importantly, the findings demonstrated that statistically significant differences emerged between expectations and evaluations, even for those classes where the actual evaluations would probably qualify as being reasonably high. In such instances, as for Classes A and C in Table 1, significant differences occurred in half of the categories. If evaluations were the only measure

for those classes in Tables 2 and 3, the lower expectations, in comparison to still lower--and significant--evaluations would surely escape notice. So, despite the fact that the students in all the social science classes under study exhibited similar directions for contextual expectations and evaluations, the actual intensity of these factors varied. In other words, making extrapolations from evaluations about the actual nature of student-teacher classroom interaction could be of questionable value.

This study also tried to relate the influence of year in college to expectations and evaluations. Here the data substantially indicated that freshmen and sophomores exhibited similar contextual expectations in all three colleges. Regarding evaluations, it was the sophomores, not the freshmen who consistently revealed higher evaluations--and in some instances statistically significant ones. This would certainly question the applicability of the findings relating situational expectations to specific classroom contexts. If a "Freshman Myth" existed for this student population, then it applied to sophomores, too. This was confirmed when freshman and sophomore expectations and evaluations were compared as reported in Table 9. Nonetheless, as the subsequent consideration of the demographic variables in Tables 10 to 16 bore out, the student populations might have been too in-group homogeneous to reveal year in college differences. This criticism would not, of course, negate the fact that within these specific classroom contexts the differences played by year in college were minimal.

Although only intended to be suggestive, consideration of the interrelationships between the CI expectations and evaluations categories with certain demographic variables, disclosed some pertinent information regarding the three colleges. In order to track down why the only significant differences developed for Gamma College were in sex and secondary class size, it was discovered that the otherwise lack in differences for expectations and evaluations might be explained by the relative homogeneity of the groups. Nevertheless, it appeared that the findings of previous evaluations studies, which noted few substantial differences for comparable demographic variables, were confirmed by this study's student population. The data also revealed that these variables failed to produce significant differences for expectations. Indeed, it seemed that greater discrepancies were exhibited with evaluations categories (see Table 15).

Another area introduced into the study for suggestive purposes sought to relate the CI expectations and evaluations categories with the six AVL value scales. Although the AVL was confirmed as to its test and retest reliability as indicated in Table 17, it did not produce any overall intercorrelation pattern. The one significant finding pointed to the need for further work relating contextual expectations with respect to student values. This result, that males at Gamma College who exhibited positive expectations for course content and requirements in relation to the AVL political scale also revealed negative expectations for concern with the instructor's style would provide a likely starting point.

If this study found that students in the three colleges responded consistently on the test and posttest of the CI expectations and evaluations instruments and the AVL value scales, then it would appear desirable to assume that further work would discover the relationships between expectations and values. A key question, left unanswered by the findings of this study, was: would the homogeneity that expressed itself in a variety of ways here, emerge in future studies of expectations and values?

Implications

This study began as an inquiry into the nature of instructional evaluations--did its findings reveal any implications for currently accepted evaluations procedures? If evaluations mean a reaction to the instructional process, then the end-of-course rating may still serve a valuable function. For as this study confirmed--in keeping with previous evaluations research--an end-of-course rating would be a reliable indicator of student reaction. But if closer student-teacher interaction becomes desirable, then, the appropriateness of expectation inputs would be obvious. Simply, presently conceived end-of-course ratings have built-in limitations. In addition, these evaluations tend to create the false impression that the students in any one class remain inert, especially if an instructor receives similar ratings over time. To sum up, even for the three relatively homogeneous student groups,

the expectations dimension should have added to an understanding of teacher-student interaction, or the lack of it.

The problem of interaction raised a question concerning the consensual expectations model. The approach called for greater participation by student and teacher in defining classroom objectives; its corollary would be the increased awareness of each other's expectations. It might be argued that in order to establish the reasonableness of this model, the study had to demonstrate the obvious. That is, student evaluations would be different from their expectations. If this was so, it resulted from the fact that end-of-course ratings failed to include the intensity of student expectations. Moreover, it seemed that student expectations continued to play a limited role in college classrooms. It became necessary, therefore, to examine whether significant differences actually developed in classrooms. Based on this study findings, the consensual approach would discourage any evaluation cast in isolation from expectations.

The implications for the study of classroom--contextual--expectations in relation to institutional--situational--expectations were most suggestive. There seemed to be differences in the role played by each set of expectations. For one, year in college--prominent in situational research--did not prove a significant factor in determining either expectations or evaluations. If the "Freshman Myth" offered by Stern meant that freshmen would be more disenchanted with their situational expectations, this study advanced that freshman and sophomore

contextual expectations and evaluations remained relatively similar. These findings suggested the need for clarification between the sets of expectations. As part of that study agenda, it would be important to learn how students report each set of expectations. Are the differences to be found in the perception of each set? The findings of this study--in keeping with Stogdill's framework--indicated that the role and performance aspects of a classroom would clearly account for some of the possible differences in reporting contextual expectations. In sum, the dynamics of classroom membership compared with institutional membership have still to be studied.

Recommendations for Further Study

This study was limited by the in-group homogeneity of the three student groups, and it might be appropriate to replicate the design in other multi-institutional settings with a greater cross-section of students. Since most evaluations research has taken place in state colleges and universities, it would be advantageous to introduce the expectations dimension into those settings, too. If single institutions must be used, then the sample group should be heterogeneous. Not only would this provide an increased awareness of classroom dynamics, but it might also encourage receptivity to greater student-teacher interaction. One recommended measure would be the Rokeach Value Scale. If this contributed to understanding the nature of student values, a more comprehensive understanding of student expectations would be in order,

too. Here, additional work in the area of instrumentation seemed appropriate. Although the present study suggested Osgood's Semantic Differential Technique to overcome the increased proliferation of instruments, perhaps other scales could be correlated with it. Any future refinements should work to validate expectations-evaluations instruments. All in all, more work needs to be done relating the nature and origin of values with expectations.

Improvements in the study design could be suggested, too. One would be to introduce certain controls into the experimental design. For example, it could prove worthwhile to manipulate the classroom situation--the independent variable--by employing the concept of consensual expectations in some groups only. Although many factors would require control, it would be important to learn the effects of this approach in actual classrooms. This type of experiment could be conducted in larger institutions with relative ease. In small colleges further controls would have to be introduced through a multi-institutional design.

Finally, further studies might consider the intervening variables of class size and other subject areas in relation to the expectations-evaluations approach suggested here. Long-range longitudinal studies would be needed to determine the impact of actually employing this approach on both students and teachers. This type of study would help to understand whether consensual expectations could become an acceptable classroom approach.

Afterword

Although events of the last ten years have added greater urgency for studies of college culture, the resulting research appeared relatively unconcerned with classroom role dynamics, least of all student contextual expectations.[1] As Twa observed in 1970: "Research and literature in the area of students' expectations are scarce . . . ."[2] Even Feldman and Newcomb--a sociologist and a social psychologist, respectively--made no attempt to distinguish between the differing research designs and definitions related to student expectations, such as the contextual or the situational ones.[3] Consequently, their findings related to classroom expectations appeared somewhat inconsistent.[4]

The research situation remained not too different from a sociologist's

---

[1]James W. Trent and Leland L. Medsker, Beyond High School (San Francisco: Jossey-Bass, Inc., 1968) and Walter L. Wallace, Student Culture: Social Structure in a Liberal Arts College (Chicago: Aldine Publishing Company, 1966).

[2]R. James Twa, Student and Instructor Expectations of Community College Instructors (Eugene, Oregon: Oregon School Study Council, 1970), pp. 6-7.

[3]Kenneth A. Feldman and Theodore M. Newcomb, The Impact of College on Students 2 vols.; (San Francisco: Jossey-Bass, Inc., 1969) I, pp. 78-82.

[4]Ibid. See also Alexander W. Astin, "The Methodology of Research on College Impact," Part I, Sociology of Education, 43 (Summer, 1970), 223-254.

observation of ten years ago:[5]

> Relatively little research data is available concerning college student and teacher role dynamics, and much of the knowledge about these roles possessed by educators and social scientists is still of an informal nature--personal, intuitive, based upon individual experience, unsystematically derived. True there is a considerable mass of literature that is tangentially related, such as the rather numerous student ratings of teachers and studies of student adjustment in college, as well as a few broad works on the academic profession. In any case, a straightforward attack on the area of student-instructor role dynamics in the college cultural setting is yet to be made. (Italics mine.)

He also noted a special reluctance of sociology to confront the classroom:[6]

> Sociologists, oddly enough, have been particularly remiss in their contributions to this area. While students have been widely utilized (indeed, sometimes exploited) as subjects for sociological research, a systematic theoretical and empirical scrutiny of the college culture and its dominant roles has thus far been avoided.

This reluctance can no longer go unnoticed.

Fortunately, a corrective was recently issued calling for a multi-disciplinary field to study parameters of classroom dynamics, appropriately named the sociology of learning.[7] If its limits were still somewhat undefined, examining the interplay of contextual and situational expectations with related sociological variables in the college setting

---

[5]Harry R. Dick, "Student-Faculty Role Consensus, Southwestern Social Science Quarterly, 41 (March, 1961), 415.

[6]Ibid., p. 416.

[7]Sarane S. Boocock, "Toward a Sociology of Learning: A Selective Review of Existing Literature," Sociology of Education, 39 (Winter, 1966), 1.

should be part of its research agenda. The concept of a sociology of learning would thereby give needed direction to classroom evaluations research, including the contextual expectations dimension.

BIBLIOGRAPHY

Allport, Gordon, Vernon, Philip E., and Lindzey, Gardner. Study of Values. Boston: Houghton Mifflin and Company, 1963.

Astin, Alexander W. "The Methodology of Research on College Impact," Part I, Sociology of Education, 43 (Summer, 1970), 223-254.

Austin, C. Grey. "The Year's Books," Journal of Higher Education, 40 (November, 1969), 589-592.

Axelrod, Joseph. Model Building for Undergraduate Colleges. U. S. Office of Education Contract No. OEC 6-10-106. Washington, D. C.: Educational Resources Information Center, 1969.

Berger, Joseph, Conner, Thomas, and McKeown, William L. "Evaluations and the Formation and Maintenance of Performance Expectations," Human Relations, 22 (December, 1969), 481-502.

Boocock, Sarance S. "Toward a Sociology of Learning: A Selective Review of Existing Literature," Sociology of Education, 39 (Winter, 1966), 1-45.

Bradshaw, P. L. and Crawford, H. L. "Perception of Characteristics of Effective University Teachers: A Scaling Analysis," Educational and Psychological Measurement, 28 (Winter, 1968), 1079-1085.

Brim, Orville G. Sociology and the Field of Education. New York: Russell Sage Foundation, 1958.

Brinton, James F. "Deriving an Attitude Scale from Semantic Differential Data," Public Opinion Quarterly, 25 (Summer, 1961), 289-295.

Buckley, Walter. Sociology and Modern Systems Theory. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1967.

Caffrey, Bernard. "Lack of Bias in Student Evaluations of Teachers," Proceedings of the 77th Annual Convention of the American Psychological Association, 1969, Vol. IV, Pt. I. Washington, D. C.: American Psychological Association, 1969.

Cartwright, Dorwin and Zander, Alvin (eds.). Group Dynamics: Research and Theory. 3d ed. New York: Harper and Row, Publishers, 1968.

Chickering, Arthur W. Education and Identity. San Francisco: Jossey-Bass, Inc., 1969.

Clark, Burton R. "Sociology of Education," Handbook of Modern Sociology, ed. Robert E. L. Faris, 734-769. Chicago: Rand McNally and Company, 1964.

Coffman, William E. "Determining Students' Concepts of Effective Teaching From Their Ratings of Instructors," Journal of Educational Psychology, 45 (May, 1954), 277-286.

Cohen, Arthur M. and Brawer, Florence B. Measuring Faculty Performance. "Monograph Series"; Washington, D. C.: American Association of Junior Colleges, 1969.

Cynamon, Manuel and Wedeen, Shirley U. "Emotional Factors in Reliability of Student Rating of Teachers," Journal of Educational Research, 51 (April, 1958), 629-632.

Dick, Harry R. "Student-Faculty Role Consensus," Southwestern Social Science Quarterly, 41 (March, 1961), 415-423.

Eble, Kenneth E. The Recognition and Evaluation of Teaching. Report of the Project to Improve College Teaching, Salt Lake City, 1970. Salt Lake City, Utah: Project to Improve College Teaching, 1970.

Eckert, Ruth E. "Ways of Evaluating Teaching," School and Society, 71 (February 4, 1950), 65-69.

Eddy, Edward D., Jr. The College Influence on Student Character. Washington, D. C.: American Council on Education, 1959.

Feldman, Kenneth A. "Studying the Impact of Colleges on Students," Sociology of Education, 42 (Summer, 1969), 207-237.

_____, and Newcomb, Theodore M. The Impact of College on Students. 2 Vols. San Francisco: Jossey-Bass, Inc., 1969.

Gage, N. L. "Appraisal of College Teaching: An Analysis of Ends and Means," Journal of Higher Education, 32 (January, 1961), 17-22.

Garverick, C. M. and Carter, H. D. "Instructors Ratings and Expected Grades," California Journal of Educational Research, 13 (November, 1962), 218-221.

Gouldner, Alvin W. The Coming Crisis in Western Sociology. New York: Basic Books, 1970.

Gowin, D. Bob and Donald E. Payne. "Evaluating Instruction: Cross Perceptions of College Students and Teachers," School Review, 70 (Spring, 1962), 207-219.

Gray, Charles E. "The Teaching Model and Evaluation of Teaching Performance," Journal of Higher Education, 40 (November, 1969), 636-642.

Gustad, John W. "Evaluation of Teaching Performance: Issues and Possibilities," Improving College Teaching, ed. Calvin B. T. Lee, 265-281. Washington, D. C.: American Council on Education, 1967.

Heath, Douglas H. Growing Up In College: Liberal Education and Maturity. San Francisco: Jossey-Bass, Inc., 1968.

Isaacson, Robert L., McKeachie, Wilbert J., and Milholland, John E. "Correlation of Teacher Personality Variables and Student Ratings," Journal of Educational Psychology, 54 (April, 1963), 110-117.

Isaacson, R. L., et al. "Dimensions of Student Evaluations of Teaching, "Journal of Educational Psychology, 55 (December, 1964), 344-351.

Katz, F. M. and Katz, Cecily, "Great Expectations," Universities Quarterly, 23 (Autumn, 1969), 420-430.

Katz, Joseph, et al. No Time For Youth. San Francisco: Jossey-Bass, Inc., 1968.

Lauterbach, C. G. and Lauterbach, D. P. "Need-Press and Expectations-Press Indices and Predictions of College Achievement," Educational and Psychological Measurement, 26 (Winter, 1966), 965-972.

Lehmann, Irvin J. and Dressel, Paul I. Critical Thinking, Attitudes and Values in Higher Education. East Lansing: Michigan State University, 1962.

Lewis, Lionel S. "Students' Images of Professors," Educational Forum, 32 (January, 1968), 185-190.

Maney, Ann C. "Authoritarianism Dimension in Student Evaluations of Faculty," Journal of Educational Psychology, 32 (January, 1959), 226-231.

McClelland, James N. "The Effect of Student Evaluations of College Instruction Upon Subsequent Evaluations," California Journal of Educational Research, 31 (March, 1970), 88-95.

McKeachie, W. J. and Soloman, D. "Student Ratings of Instructors," Journal of Educational Research, 51 (January, 1958), 379-382.

McKeachie, W. J. "Research on Teaching at the College and University Level," Handbook of Research on Teaching, ed. N. L. Gage, 1118-1172. Chicago: Rand McNally and Company, 1963.

_____. "Student Ratings of Faculty," AAUP Bulletin,55 (December, 1969), 439-444.

Megaw, Neill. "The Dynamics of Evaluation," Improving College Teaching, ed. Calvin B. T. Lee, 282-285. Washington, D. C.: American Council on Education, 1967.

Mueller, John H., Schuessler, Karl, Costner, Herbert L. Statistical Reasoning in Sociology. 2d ed. Boston: Houghton Mifflin Company, 1970.

Newcomb, Theodore M. "Student Peer-Group Influence," The American College, ed. Nevitt Sanford, 469-488. New York: John Wiley and Sons, 1962.

Oetting, S. R. "The Effect of Forcing Response on the Semantic Differential," Educational and Psychological Measurement, 67 (Autumn, 1967), 699-702.

Osgood, Charles E., Suci, George J., and Tannenbaum, Percy H. The Measurement of Meaning. Urbana, Illinois: University of Illinois Press, 1957.

Owen, Paul H. "Some Dimensions of College Teaching: An Exploratory Study Using Critical Incidents and Factor Analysis of Student Ratings." Unpublished Doctoral dissertation, University of Houston, 1967.

Perry, Richard R. "Criteria of Effective Teaching Behavior in an Institution of Higher Education." Proceedings of the Seventh Annual Forum of the Association for Institutional Research, 1967. n. p.: The Association for Institutional Research, 1967.

Pervin, Lawrence A. "Satisfaction and Perceived Self-Environment Similarity: A Semantic Differential Study of Student-College Interaction," Journal of Personality, 35 (December, 1967), 623-634.

_____. "Reality and Non-Reality in Student Expectations of College," Journal of Psychology, 64 (September, 1966), 41-48.

Phillips, Beeman N. "Individual and the Classroom Group as Frames of Reference in Determining Teacher Effectiveness," Journal of Educational Research, 58 (November, 1964), 128-131.

Plant, Walter T. and Telford, Charles W. "Changes in Personality for Groups Completing Different Amounts of College Over Two Years," Genetic Psychology Monographs, 74 (August, 1966), 3-36.

Quick, Alan F. and Wolfe, Arnold. "The Ideal Professor," Improving College and University Teaching, 13 (Summer, 1965), 133-134.

Rayder, Nicholas F. "College Student Ratings of Instructors," Journal of Experimental Education, 37 (Winter, 1968), 76-81.

Rees, Richard D. "Dimensions of Students' Points of View in Rating College Teachers," Journal of Educational Psychology, 60 (December, 1969), 476-482.

Remmers, H. H. "On Students' Perceptions of Teachers' Effectiveness," The Appraisal of Teaching in Large Universities, ed. W. J. McKeachie, 17-23. Ann Arbor: University of Michigan, 1959.

_____. Manual of Instructions for the Purdue Rating Scale for Instructors. Rev. ed. Lafayette, Indiana: Purdue University Book Store, 1960.

_____. and Drucker, A. J. "Do Alumni and Students Differ in Their Attitudes Toward Instructors?" Journal of Educational Psychology, 42 (March, 1951), 129-143.

Renner, Richard R. "A Successful Rating Scale," Improving College and University Teaching, 15 (Winter, 1967), 12-14.

Riley, John W., Jr., Ryan, Bryce F., and Lifshitz, Marcia. The Student Looks at His Teacher. New Brunswick, New Jersey: Rutgers University Press, 1950.

Rokeach, Milton. "The Role of Values in Public Opinion Research," Public Opinion Quarterly, 32 (December, 1968), 548-559.

Rudolph, Frederick. "Neglect of Students as Historical Tradition," The College and the Student, eds. L. E. Dennis and J. F. Kaufman, 47-58. Washington, D. C.: American Council on Education, 1966.

Ryans, David G. Characteristics of Teachers. Washington, D. C.: American Council on Education, 1960.

Schoen, Walter. "Faculty Evaluation, Pro and Con: The Hypocrites," Report of the Twenty-Second Annual Conference, 1969. Cazenovia, New York, New York State Association of Junior Colleges, 1969.

Simpson, Roy H. and Seidman, Jerome M. Student Evaluations of Teaching and Learning. Washington, D. C.: American Association of Colleges for Teacher Education, 1962.

Snider, James G. and Osgood, Charles E. (eds.) Semantic Differential Technique: A Sourcebook. Chicago: Aldine Publishing Company, 1969.

Stern, George C. People In Context: Measuring Person-Environment Congruence in Education and Industry. New York: John Wiley and Sons, 1970.

_____. "Myth and Reality in the American College," AAUP Bulletin, 52 (December, 1966), 408-414.

Stewart, C. T. and Malpass, L. F. "Estimates of Achievement and Ratings of Instructors," Journal of Educational Research, 59 (April, 1966), 347-350.

Stogdill, Ralph M. Individual Behavior and Group Achievement. New York: Oxford University Press, 1959.

Strand, David A. "Rationale and Instrument for Student Evaluation of Classroom Teaching," National Association of Women Deans and Counselors Journal, 30 (February, 1966), 36-39.

Taylor, Harold. Students Without Teachers: The Crisis in the University. New York: McGraw Hill Book Company, 1969.

Trent, James W. and Medsker, Leland L. Beyond High School. San Francisco: Jossey-Bass, Inc., 1968.

Turner, Richard L., et al. "How Do Student Characteristics Affect Their Evaluations of Instructors," Indiana University School of Education Bulletin, 45 (July, 1969), 46-97.

Twa, R. James. Student and Instructor Expectations of Community College Instructors. Eugene, Oregon: Oregon School Study Council, 1970.

Voeks, Virginia W. and French, G. M. "Are Student Ratings of Teachers Affected by Grades?" Journal of Higher Education, 31 (June, 1960), 330-334.

Wallace, Walter L. Student Culture: Social Structure in a Liberal Arts College. Chicago: Aldine Publishing Company, 1966.

Wedeen, Shirley U. "Comparison to Student Reaction to Similar Concurrent Teacher-Content Instruction," Journal of Educational Research, 56 (July, 1963), 540-543.

Werdell, Philip R. Course and Teacher Evaluation. 2d ed. Washington, D. C.: National Student Association, 1966.

Yamamoto, Kaoru and Dizney, H. F. "Eight Professors: A Study on College Students' Preferences Among Their Teachers," Journal of Educational Psychology, 57 (June, 1966), 146-150.

Yamamoto, Kaoru (ed.). Teaching: Essays and Readings. Boston: Houghton Mifflin Company, 1969.

APPENDIX A

Expectations Instrument

COURSE SURVEY

We are trying to learn more about students who are in college. This survey, which is being given at other colleges, seeks to learn more about what students expect to get out of their courses. To help us in our efforts we first need to know a little about you. Since you will remain strictly anonymous, we must depend upon trusting your responses.

Date of birth_____ Male_____ Female_____

Year in College: 1st___ 2nd___ 3rd___ 4th___ City or town of birth_____

City or town of current residence_____

Kindly supply the most appropriate response for each of the following.

1. What was the size of your secondary school graduating class? Circle one.
   a. less than 100 students
   b. between 100 and 200 students
   c. between 300 and 500 students
   d. more than 500 students

2. Did you graduate from a private or public secondary school? Circle one.

3. When you graduated secondary school how large was the city or town in which you lived? Circle one.
   a. less than 10,000 people
   b. between 10,000 and 50,000 people
   c. between 50,000 and 250,000 people
   d. more than 250,000 people

4. Have you ever written an evaluation of a course you took? Circle one.
   a. Yes
   b. No

5. Is this a required or elective course? Circle one.

DIRECTIONS

The purpose of this study is to determine what you expect will happen in this course. You will be asked to respond to statements that are characteristic of college courses and instruction. In order to learn more about what your expectations are for this course, you are to indicate your response for each statement on a scale.

Here is how to use the scale. After each statement you will find a scale of opposite phrases. Where you place your check-mark will depend upon which end of the scale seems more closely associated with your personal expectation.

If your expectation for a statement is one you feel is very closely related to either end of the scale, you should place your check-mark as follows:

fair__X__:____:____:____:____:____:____unfair

OR

fair____:____:____:____:____:____:__X__unfair

If your expectation for a statement is one you feel is somewhat closely related to one or the other end of the scale (but not extremely), you should place your check-mark as follows:

just____:__X__:____:____:____:____:____unjust

OR

just____:____:____:____:____:__X__:____unjust

If you feel the statement seems only slightly related to one side as opposed to the other side, then you should check as follows:

active____:____:__X__:____:____:____:____passive

OR

active____:____:____:____:__X__:____:____passive

Use the midpoint if you feel the statement is equally associated with both sides of the scale, or if the statement is unrelated to your expectation.

strong____:____:____:__X__:____:____:____weak

IMPORTANT: Place your check-marks in the middle of spaces, not on boundaries:

This

____:__X__:____:____:____:____:____.

Not This

X

PART I

1. If students request help, in what manner do you expect the instructor to respond?
   Pleasant____:____:____:____:____:____:____Unpleasant

2. Do you expect the instructor to make clear how each topic fits into the total course?
   Actively____:____:____:____:____:____:____Passively

3. What do you expect the instructor's knowledge of the subject to be?
   Superior____:____:____:____:____:____:____Inferior

4. How do you expect the instructor to react to student questions, disagreements or expression of their own ideas?
   Approving____:____:____:____:____:____:____Disapproving

5. To what extent do you expect the instructor to maintain student interest in the course?
   Low____:____:____:____:____:____:____High

6. Do you expect to find the instructor using enough examples and explanations to clarify the material?
   Sufficient____:____:____:____:____:____:____Insufficient

7. What kind of emphasis do you expect the instructor to place on the stimulation of thinking and ideas?
   Heavy____:____:____:____:____:____:____Light

8. What percentage of the students do you expect to finish the required reading prior to class?
   Small____:____:____:____:____:____:____Large

9. What kind of emphasis do you expect the instructor to place on accumulating factual information?
   Strong____:____:____:____:____:____:____Weak

10. Considering everything, how would you expect to recommend this instructor to a friend?
    Bad____:____:____:____:____:____:____Good

PART II

1. What kind of understanding of the course objectives do you expect to have?

   Hazy____:____:____:____:____:____:____Clear

2. In relation to your other courses carrying equal credit, how do you expect the amount of study and preparation time to compare?

   Just____:____:____:____:____:____:____Unjust

3. How do you expect to find the content of the assigned reading?

   Worthless____:____:____:____:____:____:____Valuable

4. What kind of challenge do you expect this course to be for you?

   Superior____:____:____:____:____:____:____Inferior

5. How do you expect to find the content of the tests and exams?

   Unfair____:____:____:____:____:____:____Fair

6. How do you expect to relate this course to your own life?

   Low____:____:____:____:____:____:____High

7. Disregarding the instructor and the way the course is taught, how would you rate the subject matter of the course?

   Valuable____:____:____:____:____:____:____Worthless

8. In thinking about your grade how do you expect it to reflect your true worth in the course?

   Fair____:____:____:____:____:____:____Unfair

9. How would you rate the ability of this survey to determine your own expectations for this course?

   Strong____:____:____:____:____:____:____Weak

10. Considering everything, how would you expect to rate this course?

    Good____:____:____:____:____:____:____Bad

APPENDIX B

Evaluations Instrument

COURSE SURVEY

We are trying to learn more about how students evaluate their courses. To help us in our efforts we first need to know a little about your. Since you will remain strictly anonymous, we must depend upon trusting your responses.

Date of birth_____Male____Female____

Year in College: 1st___2nd___3rd___4th___ City or town of birth_____

City or town of current residence_____

DIRECTIONS

The purpose of this study is to determine what you expect will happen in this course. You will be asked to respond to statements that are characteristic of college courses and instruction. In order to learn more about what your expectations are for this course, you are to indicate your response for each statement on a scale.

Here is how to use the scale. After each statement you will find a scale of opposite phrases. Where you place your check-mark will depend upon which end of the scale seems more closely associated with your personal expectation.

If your expectation for a statement is one you feel is very closely related to either end of the scale, you should place your check-mark as follows:

fair__X__:____:____:____:____:____:____unfair

OR

fair____:____:____:____:____:____:__X__unfair

If your expectation for a statement is one you feel is somewhat closely related to one or the other end of the scale (but not extremely), you should place your check-mark as follows:

just____:__X__:____:____:____:____:____unjust

OR

just____:____:____:____:____:__X__:____unjust

If you feel the statement seems only slightly related to one side as opposed to the other side, then you should check as follows:

active____:____:__X__:____:____:____:____passive

OR

active____:____:____:____:__X__:____:____passive

Use the midpoint if you feel the statement is equally associated with both sides of the scale, or if the statement is unrelated to your expectation.

strong____:____:____:____:____:____:____weak

IMPORTANT: Place your check-marks in the middle of spaces, not on the boundaries:

This — Not This

____:__X__:____:____:____:____:____. (X)

PART I

1. If students request help, in what manner does the instructor respond?
   Pleasant____:____:____:____:____:____:____Unpleasant

2. Does the instructor make clear how each topic fits into the total course?
   Actively____:____:____:____:____:____:____Passively

3. What is the instructor's knowledge of the subject?
   Superior____:____:____:____:____:____:____Inferior

4. How does the instructor react to student questions, disagreements or expression of their own ideas?
   Approving____:____:____:____:____:____:____Disapproving

5. To what extent does the instructor maintain student interest in the course:
   Low____:____:____:____:____:____:____High

6. Does the instructor use enough examples and explanations to clarify the material?
   Sufficient____:____:____:____:____:____:____Insufficient

7. What kind of emphasis does the instructor place on the stimulation of thinking and ideas?
   Heavy____:____:____:____:____:____:____Light

8. What percentage of the students finish the required reading prior to class?
   Small____:____:____:____:____:____:____Large

9. What kind of emphasis does the instructor place on accumulating factual information?
   Strong____:____:____:____:____:____:____Weak

10. Considering everything, how would you recommend this instructor to a friend?
    Bad____:____:____:____:____:____:____Good

PART II

1. What kind of understanding of the course objectives do you have?

   Hazy____:____:____:____:____:____:____Clear

2. In relation to your other courses carrying equal credit, how does the amount of study and preparation time compare?

   Just____:____:____:____:____:____:____Unjust

3. How do you find the content of the assigned reading?

   Worthless____:____:____:____:____:____:____Valuable

4. What kind of challenge is this course for you?

   Superior____:____:____:____:____:____:____Inferior

5. How do you find the content of the tests and exams?

   Unfair____:____:____:____:____:____:____Fair

6. How does this course relate to your own life?

   Low____:____:____:____:____:____:____High

7. Disregarding the instructor and the way the course is taught, how do you rate the subject matter of the course?

   Valuable____:____:____:____:____:____:____Worthless

8. In thinking about your grade how does it reflect your true worth in the course?

   Fair____:____:____:____:____:____:____Unfair

9. How do you rate the ability of this survey to determine your own evaluation of this course?

   Strong____:____:____:____:____:____:____Weak

10. Considering everything, how do you rate this course?

    Good____:____:____:____:____:____:____Bad